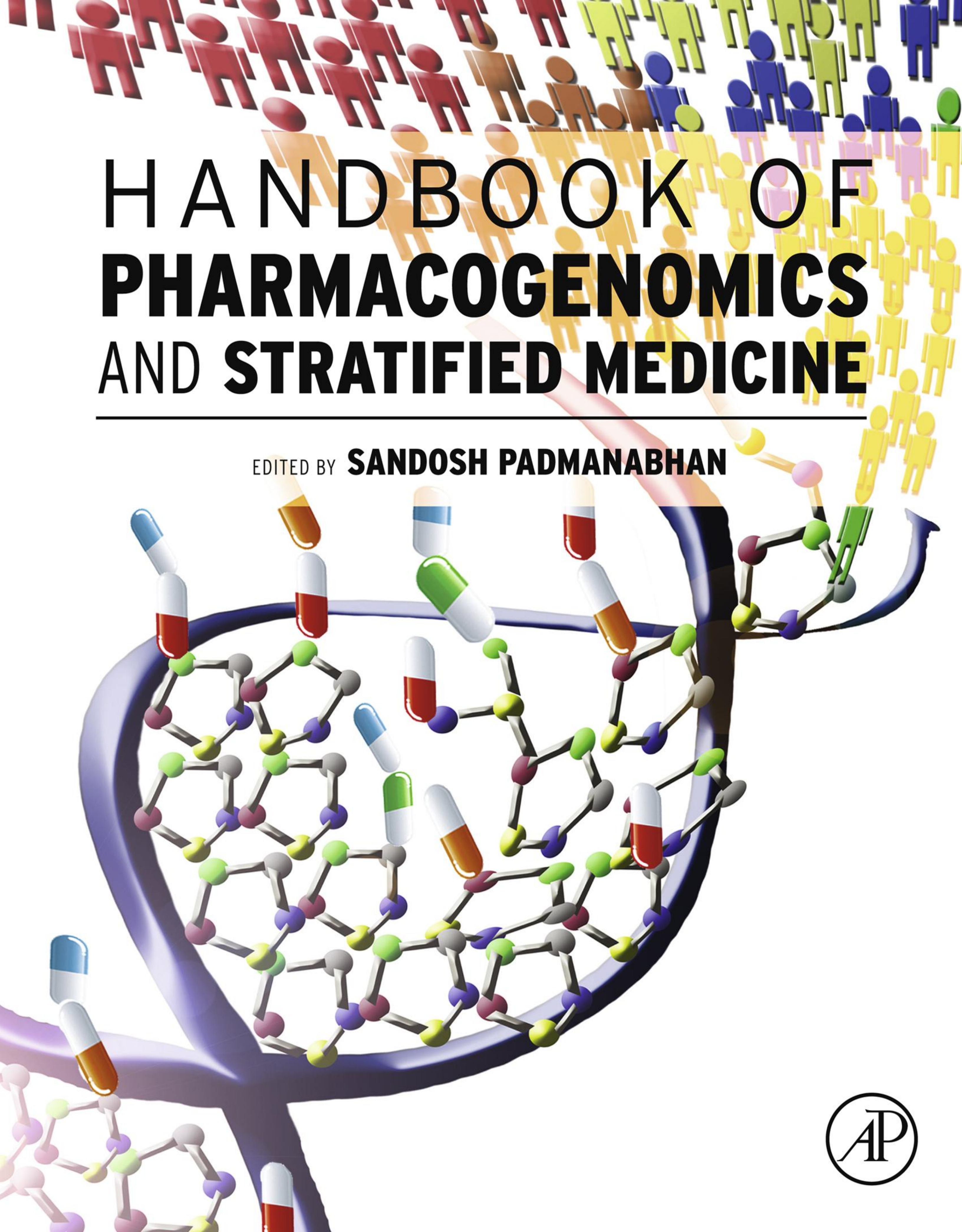
HANDBOOK OF PHARMACOGENOMICS AND STRATIFIED MEDICINE

EDITED BY **SANDOSH PADMANABHAN**



Handbook of Pharmacogenomics and Stratified Medicine

This page is intentionally left blank

Handbook of Pharmacogenomics and Stratified Medicine

Sandosh Padmanabhan
Institute of Cardiovascular and Medical Sciences,
University of Glasgow, UK

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD • PARIS
SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Copyright

Academic Press is an imprint of Elsevier

32 Jamestown Road, London NW1 7BY, UK
225 Wyman Street, Waltham, MA 02451, USA
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA

© 2014 Elsevier Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; permissions@ elsevier.com. Alternatively you can submit your request online by visiting the Elsevier website at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material.

Notices
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-386882-4

For information on all Academic Press publications
visit our website at store.elsevier.com

Typeset by SPS
www.sps.co.in

Printed and bound in the United States of America
14 15 16 17 18 10 9 8 7 6 5 4 3 2 1

ELSEVIER
Book Aid International
Working together
to grow libraries in
developing countries
www.elsevier.com • www.bookaid.org

Dedication

For Sumana and Shashank

This page is intentionally left blank

Contents

Contributors xv
Foreword xix
Preface xxi
Acknowledgements xxiii

Part I
Introduction

1. Pharmacogenomics and Stratified Medicine 3

Sandosh Padmanabhan

1.1 Overview 3
1.2 The Genetic Basis of Drug Response and Adverse Effects 5
1.3 Single-Gene Studies 5
1.4 Genome-Wide Association Studies in Pharmacogenomics 12
1.5 The Road to Personalization and Stratification 13
1.6 Pharmacogenetic Biomarkers and Companion Diagnostics 13
1.7 Economic Challenges 19
1.8 Conclusions 21
References 22

Part II
Basic Genetics

2. Basic Genetics: The Cell, Mitosis and Meiosis, and Mendelian Laws 29

Guan Wang

2.1 Overview of the Cell: Anatomy, Components, and Function 29
2.2 Cell Reproduction: Cell Cycle and Mitosis 30
2.3 Meiosis 33
2.4 Mendelian Laws 35
2.5 Public Databases for Biomedical Research in Humans 38
2.6 Conclusion 39
Glossary 39
Acknowledgments 40
References 40

3. The Human Genome, Gene Regulation, and Genomic Variation 41

Kathleen M. Murphy, Aislinn Cooper, and Edward S. Tobias

3.1 Overview: Structure of the Human Genome 41
3.2 Gene Structures 44
3.3 Gene Expression 45
3.4 Gene Regulation 46
3.5 Genomic Variation 51
3.6 Conclusion 56
References 56

4. Epigenetics 57

Tony McBryan and Peter D. Adams

4.1 Overview 57
4.2 DNA Methylation 57
4.3 Histones and Variants 61
4.4 Epigenetic Disorders and Therapies 65
4.5 Conclusion 66
References 66

Part III
Experimental and Discovery Platforms

5. Animal Models in Pharmacogenomics 73

Erin Olson and Delyth Graham

5.1 Overview 73
5.2 Rodent Models 74
5.3 Nonrodent Models of Human Disease 79
5.4 Genetic/Genomic Resources 79
5.5 Ethical Issues of Animal Experimentation 80
5.6 Caution and Limitations 80
5.7 Applications of *in vivo* Models in Pharmacogenomic Research 81
5.8 Rodent Models in Pharmacogenomics 82
5.9 Conclusion 85
References 85

6. Human Lymphoblastoid Cell Lines in Pharmacogenomics 89

Jae-Pil Jeon

6.1 Overview 89
6.2 Generation of LCLs 90
6.3 Biological Characteristics of LCLs 95
6.4 Potentials and Limitations of LCL Applications 97
6.5 Future Perspectives 105
Acknowledgments 107
References 107

7. Liver Expression Quantitative Trait Loci (eQTL) and Related Approaches in Pharmacogenomic Studies 111

Dylan M. Glubb, Amy S. Etheridge, Eric Seiser, and Federico Innocenti

7.1 Overview 111
7.2 Liver eQTL Studies 112
7.3 Future Directions 117
7.4 Conclusion 120
Acknowledgments 120
Glossary 121
Acronyms and Abbreviations 121
References 121

8. Next-Generation Sequencing 125

Pawel Herzyk

8.1 Overview 125
8.2 First-Generation Sequencing: Chain Termination Method 125
8.3 The Concept of Next-Generation Sequencing 126
8.4 Application-Independent Data Analysis 129
8.5 Genomic Applications 131
8.6 Transcriptomic Applications 133
8.7 Epigenomic Applications 135
8.8 Metagenomic Applications 137
8.9 Conclusions 141
References 141

9. Proteomics 147

Holger Husi and Amaya Albalat

9.1 Overview 147
9.2 Data Acquisition 148
9.3 Data Analysis 163
9.4 Discovery and Validation 167
9.5 Personalized Medicine 172
9.6 Conclusions 172
Acronyms 173
References 173

10. Metabolomics 181

Karl Burgess, Naomi Rankin, and Stefan Weidt

10.1 Overview 181
10.2 Metabolomics Techniques and Technologies 184
10.3 Analytical Platform 187
10.4 Analysis of Metabolomic Data 197
10.5 Conclusion 200
References 201

11. Stem Cells 207

Komal Loya

11.1 Overview 207
11.2 Stem Cell Types 208
11.3 Stem Cell Culture Methods 211
11.4 Molecular Signatures of Stem Cells 212
11.5 Epigenomics of Stem Cells 215
11.6 Pharmacogenomics of Stem Cells 218
11.7 Potential of Stem Cells 219
11.8 Future Prospects 226
Abbreviations 227
References 228

Part IV

Statistics, Bioinformatics and Databases

12. Fundamentals of Complex Trait Genetics and Association Studies 235

Jahad Alghamdi and Sandosh Padmanabhan

12.1 Overview 235
12.2 The Phenotype 235
12.3 Population Genetics 239
12.4 Association Studies 247
12.5 Conclusions 256
References 256

13. Bioinformatics: Concepts, Methods, and Data 259

Scott W. Robinson, Avid M. Afzal, and David P. Leader

13.1 Overview 259
13.2 Analysis of Sequences of Nucleic Acids and Proteins 260
13.3 Molecular Evolution and Phylogenetics 264
13.4 Molecular Graphics and Protein Modeling 269
13.5 Bioinformatics Databases 276
Glossary 284
References 284

14. PharmGKB: The Pharmacogenomics Knowledgebase 289

Julia M. Barbarino, Michelle Whirl-Carrillo, Teri E. Klein, and Russ B. Altman

14.1 Overview 289
14.2 The PharmGKB Website 290
14.3 Variant Annotations 291
14.4 Clinical Annotations 294
14.5 VIPs and Pathways 297
14.6 Clinical Implementation: Consortia, CPIC, and Dosing Guidelines 301
14.7 Projects in Development 303
14.8 Conclusion 304
Glossary 305
Acronyms and Abbreviations 305
Acknowledgments 306
References 306

Part V

Clinical Trials

15. Clinical Trials in Pharmacogenomics and Stratified Medicine 309

Sandosh Padmanabhan

15.1 Overview 309
15.2 Phase I Studies 309
15.3 Phase II Studies 312
15.4 Phase III Studies 314
15.5 Phase IV Trials 319
15.6 Retrospective Analysis of Clinical Datasets 319
References 319

Part VI

Fundamental Pharmacogenomics

16. CYP450 and Ethnicity 323

Joseph McGraw

16.1 Overview 323
16.2 Variation: Importance of Race/Ethnicity 326
16.3 Variation: CYP450 Phenotyping 327
16.4 CYP450 Variation in Different Racial/Ethnic Populations—CYP450 Families 329
16.5 Future Perspectives 337
References 337

17. Pharmacokinetic Pharmacogenomics 341

Safaa Mohammed M. Alsanosi, Craig Skiffington, and Sandosh Padmanabhan

17.1 Overview 341
17.2 Principles of Pharmacokinetics 341
17.3 ADME: Pharmacogenomics 349
17.4 Conclusions 360
References 361

18. Pharmacodynamic Pharmacogenomics 365

Linsay McCallum, Stefanie Lip, and Sandosh Padmanabhan

18.1 Overview 365
18.2 Basic Pharmacodynamics 365
18.3 Pharmacodynamic Pharmacogenetics 373
18.4 Conclusion: Pharmacodynamic Pharmacogenomics into Clinical Practice 379
Glossary 380
Acronyms and Abbreviations 381
References 381

19. MicroRNA: Regulation of P450 and Pharmacogenetics 385

Miki Nakajima and Tsuyoshi Yokoi

19.1 Overview 385
19.2 miRNA Biogenesis 385
19.3 miRNA Function 386
19.4 miRNA Target Identification 386
19.5 Regulation of Drug-Metabolizing P450s 389
19.6 Regulation of Cancer-Related P450s 391
19.7 Modulation of miRNA Expression and its Toxicological Outcome 394
19.8 miRNA Pharmacogenetics Associated with Disease or Drug Response 395
19.9 miRNAs as Biomarkers and Therapeutic Targets 397
19.10 miRNAs and Drug Response 397
19.11 Conclusions 397
Glossary 398
Abbreviations 398
References 398

Part VII
Application in Therapeutics

20. Adverse Drug Reactions 405

Lauren Walker, Vincent Yip, and Munir Pirmohamed

20.1 Overview 405
20.2 Phase I Metabolism 408
20.3 Phase II Metabolism 410
20.4 Type B Adverse Drug Reactions 414
20.5 Conclusion 424
Abbreviations and Acronyms 425
References 426

21. HLA and the Pharmacogenomics of Drug Hypersensitivity 437

Katie D. White, Silvana Gaudieri, and Elizabeth J. Phillips

21.1 Overview 437
21.2 Introduction 437
21.3 Classification of Drug Hypersensitivity Syndromes 440
21.4 Immunopathogenesis of Drug Hypersensitivity Syndromes 442
21.5 HLA and Hypersensitivity Reactions 444
21.6 Stratified Approaches to Prevent Drug Hypersensitivity and their Translation into Clinical Medicine 455
21.7 The Future Direction of Stratified Medicine in Drug-Induced Hypersensitivity 458
Acronyms 459
References 459

22. Drug-Induced Liver Injury 467

Thomas J. Urban and Hans Tillmann

22.1 Overview 467
22.2 Epidemiology and Public Health Impact 467
22.3 Etiology 468
22.4 Pharmacogenomics and Stratified Medicine in DILI 469
22.5 Difficulties in Identifying Biomarkers for DILI 470
22.6 DILI-Associated Genetic Variants 471
22.7 Utility of Dili Risk Variants in the Clinic 472
22.8 Future Directions 473
22.9 Conclusions 475
References 475

23. Hemostasis and Anticoagulants 479

Matthew R. Wilson and R. Campbell Tait

23.1 Overview 479
23.2 Hemostasis 479
23.3 Indications for Anticoagulation 482
23.4 Parenteral Anticoagulants 484
23.5 Oral Anticoagulants 488
23.6 Conclusion and Directions for Future Research 494
References 494

24. Pharmacogenomics of Warfarin 497

Ann K. Daly

24.1 Overview 497
24.2 Pharmacology of Warfarin 498
24.3 Nongenetic Factors Affecting Warfarin Dose Requirement 499
24.4 Genes with Major Effects on Warfarin Dose Requirement 500
24.5 Additional Pharmacogenomic Factors Affecting Warfarin Dosage 502
24.6 Application of Pharmacogenomics to Warfarin Prescription 503
24.7 Conclusion and Future Research Directions 504
References 504

25. Pharmacogenomics of Clopidogrel 509

Betti Giusti, Claudia Saracini, Silvia Galora, and Rossella Marcucci

25.1 Overview 509
25.2 Clopidogrel Response Variation 514
25.3 Genetic Dissection of Clopidogrel Response 518
25.4 Current Status 527
25.5 Conclusions and Future Directions 535
Abbreviations 536
References 536

26. Pharmacogenetics of Antidepressant Drugs 543

Chiara Fabbri, Alessandro Minarini, Yoshihiko Matsumoto, and Alessandro Serretti

26.1 Overview 543
26.2 Candidate Gene Studies 544
26.3 Genome-Wide Approaches 553
26.4 Conclusion 555
Key Concepts 556
Glossary 556
Acronyms and Abbreviations 556
References 557

27. Personalized Medicine of Alzheimer's Disease 563

Ramón Cacabelos, Pablo Cacabelos, and Clara Torrellas

27.1 Overview 563
27.2 Toward a Personalized Medicine for Dementia and Neurodegenerative Disorders 564
27.3 Genomics of Alzheimer's Disease 565
27.4 Pathogenic Events 565
27.5 Biomarkers and Comorbidity 569
27.6 Therapeutic Strategies 571
27.7 Pharmacogenomics 573
27.8 Future Perspective 584
27.9 Conclusion 585
Appendix A 586
Appendix B 608
References 611

28. Asthma 617

Dominick Shaw, Michael Portelli, and Ian Sayers

28.1 Overview 617
28.2 Clinical Aspects 618
28.3 Asthma Pathophysiology 623
28.4 Current Management of Asthma 629
28.5 Stratified Medicine for Asthma 631
28.6 Drug Development in Asthma and Emerging Targets 645
28.7 Conclusions and Future Directions 647
Acknowledgments 647
References 647

29. Rheumatoid Arthritis 653

Jisna R Paul, Deepali Sen, and Prabha Ranganathan

29.1 Overview 653
29.2 Pharmacogenetics of Nonbiologic DMARDs 655
29.3 Pharmacogenetics of Biologic DMARDS 668
29.4 Conclusions and Future Directions 674
References 674

30. Pharmacogenomics of Oral Antidiabetic Drugs 683

Vangelis G. Manolopoulos and Georgia Ragia

30.1 Overview 683
30.2 Genetic Basis of Diabetes 683
30.3 Management of Diabetes 684
30.4 The Need for Pharmacogenomics and Stratified Medicine 685
30.5 Genetic Variants Associated with Response to Oral Hypoglycemic Agents 685
30.6 Stratified Oral Antidiabetic Medicine: Pharmacogenomics Application 709
30.7 Conclusions: Future Perspectives 710
Abbreviations 710
References 710

31. Genomics and Pharmacogenomics of Lipid-Lowering Therapies 715

Kieran F. Docherty and Sandosh Padmanabhan

31.1 Overview 715
31.2 Lipid Metabolism 716
31.3 Lipids and CVD Risk, Epidemiology, and Public Health Impact 718
31.4 Genetic Basis of Lipid Disorders and Lipid Levels 722
31.5 Conclusion and Future Directions 741
References 742

32. Hypertension Pharmacogenomics 747

Yan Gong, Caitrin W. McDonough, Sandosh Padmanabhan, and Julie A Johnson

32.1 Overview 747
32.2 Epidemiology and Public Health Impact 747
32.3 Genetic Basis of Hypertension 749
32.4 Need for Stratified Approaches to Treatment with Antihypertensives 754
32.5 Pharmacogenomics of Blood Pressure Response to Antihypertensive Drugs 758
32.6 Pharmacogenomics of Antihypertensive Treatment—Related Outcomes 764
32.7 Future Perspectives 771
References 772

33. QTc and Sudden Cardiac Death 779

Catriona Brown and Sandosh Padmanabhan

33.1 Overview 779
33.2 Genetic Syndromes 784
33.3 Drug-Induced Variation in QTc 787
33.4 Public Health Impact 796
33.5 QTc in New Drug Discovery and Drug Safety 797
33.6 Pharmacogenomic Studies 798
33.7 Future Directions 801
Glossary 802
Acronyms and Abbreviations 803
References 803
Suggested Websites 806

34. Stratified Medicine for Pancreatic Cancer 807

David K. Chang, Nigel B. Jamieson, Sean M. Grimmond, and Andrew V. Biankin

34.1 Overview 807
34.2 Current Treatment Paradigm 807
34.3 Stratified Therapeutic Approach 808
34.4 Preclinical Models of Pancreatic Cancer 812
34.5 Central Knowledge Database 812
34.6 Biorepository and Privacy Issues 813
34.7 Conclusion 813
References 813

35. Pharmacogenomics in Anesthesia 815

Shiv Kumar Singh and Tushar Dixit

35.1 Overview 815
35.2 Pharmacokinetic Pharmacogenetics In Anesthesia 815
35.3 Pharmacodynamics 820
35.4 Conclusion 830
References 830

36. Tuberculosis 835

Geetha Ramachandran and Soumya Swaminathan

36.1 Overview 835
36.2 TB Etiology and Natural History 835
36.3 Burden of TB 836
36.4 Burden of TB Among Children 838
36.5 TB Diagnosis 838
36.6 Chemotherapy for TB 840
36.7 Risk Factors: Genetics, HIV, Diabetes, and Smoking 842
36.8 TB in Children 844
36.9 Pharmacogenomics of TB 845
36.10 Pharmacogenomics and TB 845
36.11 Conclusion 853
References 853

37. Hepatitis C Virus 859

A.J. Thompson, P. Clark, and A. Aghemo

37.1 Overview 859
37.2 *IL28B* Polymorphism and HCV Outcomes 860
37.3 *ITPA* Polymorphisms and Ribavirin-Induced Hemolytic Anemia 864
37.4 Genetic Determinants of Liver Fibrosis Progression 865
37.5 Clinical Translation 866
37.6 Conclusion 872
Acknowledgments 872
Disclosures 872
Abbreviations 872
References 873

38. Pharmacogenomics of Antifungal Agents 879

H.R. Ashbee and M.H. Gilleece

38.1 Overview of Antifungal Therapy 879
38.2 Public Health Impact of Fungal Diseases 886
38.3 The Need for Pharmacogenomics and Stratified Medicine 886
38.4 Genetic Variability Associated with Antifungal Drugs 887
38.5 Current Successes in Stratified Medicine 895
38.6 Future Perspectives 896
References 896

39. Advances in Understanding and Treatment of Human African Trypanosomiasis: Divergent Diseases Caused by Distinct Parasites 901

Federica Giordani, Roy Mwenechanya, and Michael P. Barrett

39.1 Overview 901
39.2 The Genetics of Trypanosome Lytic Factors and Parasite Factors Enabling Protection Against Lysis 903
39.3 Immune Responses and Immunoavoidance 905
39.4 Pathologies Associated with Hat 906
39.5 Diagnosis of Hat 907
39.6 Staging the Disease 908
39.7 Treatment 909
39.8 Conclusions and Future Perspectives 913
Glossary 914
References 914

Part VIII

Implementation and Regulatory

40. Implementing Clinical Pharmacogenetics: Point-of-Care and Pre-Emptive Testing 921

Stuart A. Scott

40.1 Overview 921
40.2 Clinical Pharmacogenetics and Implementation 922
40.3 Point-of-Care Pharmacogenetic Testing 924
40.4 Pre-Emptive Pharmacogenetic Testing 925
40.5 Conclusion and Future Perspectives 927
References 928

41. Ethical Considerations in Pharmacogenomic Testing and Research in Pediatrics **931**

Chi Dang Hornik, Cait O'Donnell, and Raymond Barfield

41.1 Overview 931
41.2 Historical Perspective 933
41.3 Ethical Significance of Benefit, Risk, and Uncertainty in Pediatric Pharmacogenomics 935
41.4 Federal Regulations and National Healthcare Priorities Related to Pharmacogenomics 943
41.5 Translation of Pharmacogenomics into the Clinical Setting 945
41.6 Future Implications of Pharmacogenomics Testing in Children 946
41.7 Conclusion 948
References 948

42. PGx/Biomarker Utilization for Regulatory Decision Making **951**

Yu Sakamoto, Yasuto Otsubo, Akihiro Ishiguro, and Yoshiaki Uyama

42.1 Introduction 951
42.2 Regulatory Activities and Guidelines Relating to PGx/Biomarker 951
42.3 Biomarker Qualification 956
42.4 Clinical Trial Design and Drug Development Utilizing PGx/Biomarker 960
42.5 Future Challenges 963
Disclaimer 964
Glossary 964
Acronyms and Abbreviations 964
References 965

Part IX

Global Perspective

43. Population Diversity and Pharmacogenomics in Africa **971**

Collen Masimirembwa, Collet Dandara, and Julia Hasler

43.1 Overview 971
43.2 Population Diversity in Africa 973
43.3 Population Genetics, Diversity, Ancestry, and Admixture 979
43.4 Pharmacogenomics in Africa 985
43.5 Future Perspectives 992
43.6 Conclusions 994
Key Concepts 994
Acknowledgments 995
Glossary 995
Acronyms and Abbreviations 995
References 996

44. Pharmacogenomics in China **999**

Wei Zhang, Hong-Hao Zhou, Zhao-Qian Liu, Xiao-Ping Chen, and Yi-Jing He

44.1 Overview 999
44.2 Pharmacogenomics of Drug-Metabolizing Enzymes in Chinese Populations 999
44.3 Pharmacogenomics of Drug Transporters in Chinese Populations 1002
44.4 Pharmacogenomics of Drug Receptors in Chinese Populations 1006
44.5 Pharmacogenomics of ION Channels in Chinese Populations 1009
44.6 Environmental Factors in Drug-Metabolizing Enzyme and Transporter Activity 1011
44.7 Translational Approaches to Pharmacogenomics in China 1012
44.8 Conclusion and Future Perspectives 1012
Acknowledgments 1012
References 1012

45. Pharmacogenomics in Brazil **1015**

Guilherme Suarez-Kurtz and Mara Helena Hutz

45.1 Overview 1015
45.2 The Brazilian Pharmacogenetics Network–Refargen 1015
45.3 PGX of Immunosuppressants in Renal Transplant Patients 1028
45.4 PGX of Antipsychotics in Schizophrenic Patients 1030
45.5 Conclusions 1032
Acknowledgments 1033
References 1033

46. Pharmacogenomics in India **1037**

Umamaheswaran Gurusamy and Deepak Gopal Shewade

46.1 Overview 1037
46.2 Indian Population Structure and Diversity 1038
46.3 Frequency Distribution and Pattern of Genetic Variants of Pharmacogenes 1039
46.4 Pharmacogenetic Studies in India 1052
46.5 Conclusions and Future Directions 1056
References 1057

Glossary 1061
Acronyms and Abbreviations 1065
Index 1071

This page is intentionally left blank

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Peter D. Adams (57), Institute of Cancer Sciences, University of Glasgow and Beatson Institute for Cancer Research, Glasgow, UK

Avid M. Afzal (259), Unilever Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, UK

A. Aghemo (857), A.M. and A. Migliavacca Center for Liver Disease, First Division of Gastroenterology, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Università degli Studi di Milano, Italy

Amaya Albalat (147), School of Natural Sciences, University of Stirling, UK

Jahad Alghamdi (235), Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Safaa Mohammed M. Alsanosi (341), BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Russ B. Altman (289), Department of Genetics, Department of Bioengineering, Stanford University, USA

H.R. Ashbee (877), Mycology Reference Centre, Department of Microbiology, Leeds Teaching Hospitals NHS Trust, UK

Julia M. Barbarino (289), Department of Genetics, Stanford University, USA

Raymond Barfield (929), Pediatric Quality of Life Program, Duke University, USA

Michael P. Barrett (899), Wellcome Trust Centre for Molecular Parasitology and Glasgow Polyomics, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Andrew V. Biankin (807), Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, UK

The Kinghorn Cancer Centre, Cancer Division, Garvan Institute of Medical Research, Australia

St Vincent's Clinical School, Faculty of Medicine, University of NSW, Australia

Department of Surgery, Bankstown Hospital, Australia

South Western Sydney Clinical School, Faculty of Medicine, University of NSW, Australia

Catriona Brown (779), Institute of Cardiovascular and Medical Sciences, Glasgow, UK

Karl Burgess (181), Glasgow Polyomics, University of Glasgow, UK

Institute of Infection, Immunity and Inflammation, University of Glasgow, UK

Pablo Cacabelos (563), Chair of Genomic Medicine, Camilo José Cela University, Madrid, Spain; EuroEspes Biomedical Research Center, Corunna, Spain

Ramón Cacabelos (563), Chair of Genomic Medicine, Camilo José Cela University, Madrid, Spain; EuroEspes Biomedical Research Center, Corunna, Spain

David K. Chang (807), Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, UK

The Kinghorn Cancer Centre, Cancer Division, Garvan Institute of Medical Research, Australia

St Vincent's Clinical School, Faculty of Medicine, University of NSW, Australia

Department of Surgery, Bankstown Hospital, Australia

South Western Sydney Clinical School, Faculty of Medicine, University of NSW, Australia

Xiao-Ping Chen (997), Pharmacogenetics Research Institute, Institute of Clinical Pharmacology, Central South University, PR China

P. Clark (857), Department of Gastroenterology, Princess Alexandra Hospital and School of Medicine, University of Queensland, Australia

Aislinn Cooper (41), Medical Genetics, School of Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Ann K. Daly (497), Institute of Cellular Medicine, Newcastle University, UK

Collet Dandara (969), Department of Clinical Laboratory Science, University of Cape Town, South Africa

Tushar Dixit (815), Mersey Deanery, UK

Kieran F. Docherty (715), BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, UK

Amy S. Etheridge (111), Eshelman School of Pharmacy, Institute for Pharmacogenomics and Individualized Therapy, University of North Carolina at Chapel Hill, USA

Chiara Fabbri (543), Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy

Silvia Galora (509), Department of Experimental and Clinical Medicine, University of Florence, Atherothrombotic Diseases Center, Careggi Hospital, Italy

Silvana Gaudieri (437), School of Anatomy, Physiology and Human Biology, University of Western Australia, Australia

Institute for Immunology and Infectious Diseases, Murdoch University, Australia

M.H. Gilleece (877), Department of Haematology, St James' Institute of Oncology, Leeds Teaching Hospitals NHS Trust, UK

Federica Giordani (899), Wellcome Trust Centre for Molecular Parasitology and Glasgow Polyomics, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Betti Giusti (509), Department of Experimental and Clinical Medicine, University of Florence, Atherothrombotic Diseases Center, Careggi Hospital, Italy

Dylan M. Glubb (111), Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Australia

Yan Gong (747), Department of Pharmacotherapy and Translational Research, Center for Pharmacogenomics, University of Florida, USA

Delyth Graham (73), BHF Glasgow Cardiovascular Research Centre, University of Glasgow, UK

Sean M. Grimmond (807), Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, UK

Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, University of Queensland, Australia

Umamaheswaran Gurusamy (1035), Pharmacogenomics Laboratory, Department of Pharmacology, Jawaharlal Institute of Postgraduate Medical Education and Research, India

Julia Hasler (969), African Institute of Biomedical Science and Technology, Zimbabwe

Yi-Jing He (997), Pharmacogenetics Research Institute, Institute of Clinical Pharmacology, Central South University, PR China

Pawel Herzyk (125), Glasgow Polyomics and The Institute of Molecular Cell and Systems Biology, University of Glasgow, UK

Chi Dang Hornik (929), Neonatal Intensive Care, Duke Children's Hospital & Health Center, Duke University Medical Center, USA

Holger Husi (147), BHF Glasgow Cardiovascular Research Centre, ICAMS University of Glasgow, UK

Mara Helena Hutz (1013), Department of Genetics, Universidade Federal do Rio Grande do Sul, Brazil

Federico Innocenti (111), Eshelman School of Pharmacy, Institute for Pharmacogenomics and Individualized Therapy, University of North Carolina at Chapel Hill, USA

Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, USA

School of Medicine, University of North Carolina at Chapel Hill, USA

Akihiro Ishiguro (949), PMDA Omics Project, Pharmaceuticals & Medical Devices Agency (PMDA), Japan

Nigel B. Jamieson (807), Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, UK,

Academic Unit of Surgery, Faculty of Medicine, Glasgow Royal Infirmary, University of Glasgow, UK

Jae-Pil Jeon (89), Division of Brain Diseases, Center for Biomedical Science, Korea National Institute of Health, Korea Centers for Disease Control and Prevention, Republic of Korea

Julie A. Johnson (747), Department of Pharmacotherapy and Translational Research, Center for Pharmacogenomics, University of Florida, USA

Department of Medicine, University of Florida, USA

Teri E. Klein (289), Department of Genetics, Stanford University, USA

David P. Leader (259), School of Life Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Stefanie Lip (365), BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Zhao-Qian Liu (997), Pharmacogenetics Research Institute, Institute of Clinical Pharmacology, Central South University, PR China

Komal Loya (207), College of Medical Veterinary and Life Sciences, Institute of Cardiovascular and Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, UK

Vangelis G. Manolopoulos (683), Laboratory of Pharmacology and Clinical Pharmacology, Medical School, Democritus University of Thrace, Alexandroupolis, Greece

Rossella Marcucci (509), Department of Experimental and Clinical Medicine, University of Florence, Atherothrombotic Diseases Center, Careggi Hospital, Italy

Collen Masimirembwa (969), African Institute of Biomedical Science and Technology, Zimbabwe

Yoshihiko Matsumoto (543), Department of Psychiatry, School of Medicine, Yamagata University, Japan

Tony McBryan (57), Institute of Cancer Sciences, University of Glasgow and Beatson Institute for Cancer Research, Glasgow, UK

Linsay McCallum (365), BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Caitrin W. McDonough (747), Department of Pharmacotherapy and Translational Research, Center for Pharmacogenomics, University of Florida, USA

Joseph McGraw (323), Metabolism Laboratory, Department of Pharmaceutical and Administrative Sciences, Concordia University School of Pharmacy, USA

Alessandro Minarini (543), Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy

Kathleen M. Murphy (41), Medical Genetics, School of Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Roy Mwenechanya (899), Wellcome Trust Centre for Molecular Parasitology and Glasgow Polyomics, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Miki Nakajima (385), Drug Metabolism and Toxicology, Faculty of Pharmaceutical Sciences, Kanazawa University, Japan

Cait O'Donnell (929), Ethics in Military Medicine Research Group, McMaster University, UK

Erin Olson (73), BHF Glasgow Cardiovascular Research Centre, University of Glasgow, UK

Yasuto Otsubo (949), PMDA Omics Project, Pharmaceuticals & Medical Devices Agency (PMDA), Japan

Sandosh Padmanabhan (3), (235), (309), (341), (365), (715), (747), Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Jisna R Paul (653), Division of Rheumatology, Department of Medicine, Washington University School of Medicine, St. Louis, USA

Elizabeth J. Phillips (437), Department of Medicine, Vanderbilt University Medical Center, USA

Institute for Immunology and Infectious Diseases, Murdoch University, Australia

Munir Pirmohamed (405), The Wolfson Centre for Personalized Medicine, Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool, UK

Michael Portelli (617), Division of Respiratory Medicine, University of Nottingham, UK

Georgia Ragia (683), Laboratory of Pharmacology and Clinical Pharmacology, Medical School, Democritus University of Thrace, Alexandroupolis, Greece

Geetha Ramachandran (835), National Institute for Research in Tuberculosis, Indian Council of Medical Research, India

Prabha Ranganathan (653), Division of Rheumatology, Department of Medicine, Washington University School of Medicine, St. Louis, USA

Naomi Rankin (181), Glasgow Polyomics, University of Glasgow, UK

Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Scott W. Robinson (259), Institute of Cardiovascular and Medical Sciences, BHF Cardiovascular Research Centre, University of Glasgow, UK

Yu Sakamoto (949), PMDA Omics Project, Pharmaceuticals & Medical Devices Agency (PMDA), Japan

Claudia Saracini (509), Department of Experimental and Clinical Medicine, University of Florence, Atherothrombotic Diseases Center, Careggi Hospital, Italy

Ian Sayers (617), Division of Respiratory Medicine, University of Nottingham, UK

Stuart A. Scott (919), Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, USA

Eric Seiser (111), Eshelman School of Pharmacy, Institute for Pharmacogenomics and Individualized Therapy, University of North Carolina at Chapel Hill, USA

Deepali Sen (653), Division of Rheumatology, Department of Medicine, Washington University School of Medicine, St. Louis, USA

Alessandro Serretti (543), Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy

Dominick Shaw (617), Division of Respiratory Medicine, University of Nottingham, UK

Deepak Gopal Shewade (1035), Pharmacogenomics Laboratory, Department of Pharmacology, Jawaharlal Institute of Postgraduate Medical Education and Research, India

Shiv Kumar Singh (815), Royal Liverpool University Hospitals, UK

Craig Skiffington (341), BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Guilherme Suarez-Kurtz (1013), Pharmacology Division, Instituto Nacional de Câncer, Brazil

Soumya Swaminathan (835), National Institute for Research in Tuberculosis, Indian Council of Medical Research, India

R. Campbell Tait (479), Department of Haematology, Glasgow Royal Infirmary, UK

A.J. Thompson (857), Department of Gastroenterology, St Vincent's Hospital, University of Melbourne, Victorian Infectious Diseases Reference Laboratory, Australia

Hans Tillmann (467), Department of Medicine, Division of Gastroenterology, Duke University Medical Center, Durham, NC, USA

Edward S. Tobias (41), Medical Genetics, School of Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Clara Torrellas (563), Chair of Genomic Medicine, Camilo José Cela University, Madrid, Spain; EuroEspes Biomedical Research Center, Corunna, Spain

Thomas J. Urban (467), Center for Human Genome Variation, USA

Yoshiaki Uyama (949), PMDA Omics Project, Pharmaceuticals & Medical Devices Agency (PMDA), Japan

Department of Regulatory Science of Medicine, Graduate School of Medicine, Chiba University, Japan

Lauren Walker (405), The Wolfson Centre for Personalized Medicine, Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool, UK

Guan Wang (29), Institute of Cardiovascular and Medical Sciences, College of Medicine, Veterinary and Life Sciences, University of Glasgow, UK

Stefan Weidt (181), Glasgow Polyomics, University of Glasgow, UK

Institute of Infection, Immunity and Inflammation, University of Glasgow, UK

Michelle Whirl-Carrillo (289), Department of Genetics, Stanford University, USA

Katie D. White (437), Department of Medicine, Vanderbilt University Medical Center, USA

Matthew R. Wilson (479), Department of Haematology, Glasgow Royal Infirmary, UK

Vincent Yip (405), The Wolfson Centre for Personalized Medicine, Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool, UK

Tsuyoshi Yokoi (385), Drug Metabolism and Toxicology, Faculty of Pharmaceutical Sciences, Kanazawa University, Japan

Wei Zhang (997), Pharmacogenetics Research Institute, Institute of Clinical Pharmacology, Central South University, PR China

Hong-Hao Zhou (997), Pharmacogenetics Research Institute, Institute of Clinical Pharmacology, Central South University, PR China

Foreword

The *Handbook of Pharmacogenomics and Stratified Medicine*, edited by Sandosh Padmanabhan, has been designed to bring clinicians and biomedical scientists up to date with the quickly expanding field of pharmacogenomics, which is rapidly moving into everyday clinical practice. All 46 well-crafted chapters, including several consisting of practical clinical examples, provide an excellent grounding in pharmacogenomics in all its aspects.

Clear and well-designed figures and tables take us on an exciting journey from the empirical medicine of the past through the development of drug response biomarkers and so-called companion diagnostics to the stratified medicine that we either increasingly practice or will be practicing in the near future.

The timing of the publication of this handbook is excellent, as it follows and expands on the recent report on stratified medicine published by the Academy of Medical Sciences as well as the increasing consensus among major pharmaceutical companies that the momentum is there for stratified medicine not only in cancer but in several other common diseases of our civilization. Publication also coincides with several public-private innovation initiatives in stratified medicine, including the Stratified Medicine Scotland—Innovation Centre and the consultation led by the UK Technology Strategy Board on the Catapult Related to Stratified Medicine and Diagnostics.

If you, as a clinician or scientist, hope to become a part of the pharmacogenomic revolution in healthcare, the *Handbook of Pharmacogenomics and Stratified Medicine* is for you.

Anna Dominiczak

Regius professor of medicine,
vice-principal and head of the College of Medical,
Veterinary and Life Sciences,
University of Glasgow, UK
Glasgow, 28 August 2013

This page is intentionally left blank

Preface

The last decade has seen an exponential growth of genomics in terms of technological innovation and the mapping of DNA sequence variants to a multitude of traits. At the same time, there has been a steady increase in the number of drug labels that have been updated with relevant US FDA pharmacogenetic information and recommendations. Moreover, the pharmaceutical industry is actively leveraging pharmacogenomics to achieve more effective and efficient clinical drug development. The success of companion diagnostics in cancer therapy has provided a strategic direction for rapid progress in the field, with the European Medicines Agency and the US Food and Drug Administration now providing explicit guidance to industry and researchers.

Escalating investment in pharmacogenomics by industry, academia, government, and institutions reflects recognition of the discipline's value. Pharmacogenomics' evolving maturity is highlighted by pragmatic terminology such as *stratified medicine*, which "reflects the realistic effects of medicines at population level." This distancing from the term *personalized medicine* "reflects the possibly over-ambitious promise of individualized unique drug targeting and development." There is now a need for a well-organized technology infrastructure, professional training, and an internationally aligned ethical, legal, and regulatory framework to transform discoveries into clinically useful applications. A rapid transition from hypothesis to clinical application requires a multidimensional effort involving basic genetics, pharmacology, clinical trials, biostatistics, epidemiology, and ethical and regulatory issues.

The goal of *Handbook of Pharmacogenomics and Stratified Medicine* is to be accessible to readers from all of the disciplines just listed by providing a detailed overview of pharmacogenomics and stratified medicine. The first half of the book covers the principles of genetics, analysis, technology, and pharmacology, which together form the foundation for the second half focussing on disease domains. Each disease specific chapter first provide background on disease, epidemiology, and treatment before turning to the relevant pharmacogenomics.

It is envisaged that this book will help interested clinicians and nonclinicians obtain a broad and reasonably in-depth understanding of the field of pharmacogenomics and, in so doing, enhance research and quickly move discoveries into clinical implementation.

Sandosh Padmanabhan

This page is intentionally left blank

Acknowledgements

The development and completion of this book would not have been possible without the help and support of some very talented and committed individuals. Firstly, I would like to thank Christine Minihane, who first discussed with me the idea of developing a book that would help a wide spectrum of readers who are interested in the growing field of pharmacogenomics. Christine's successor Graham Nisbet continued supporting development of this book and without his encouragement, this book would not have been completed. I am very grateful to the excellent support from Catherine A Van Der Laan at Elsevier during all the stages of the project.

I would also like to express my thanks to all of the book's contributors. Many of them have busy schedules filled with clinical and research commitments. It is greatly appreciated that, because of them, all of the chapters are of high quality and will enhance the value of the Handbook of Pharmacogenomics and Stratified Medicine to its readers. In addition, I am grateful to Professor Anna F Dominiczak for writing the foreword.

A team of undergraduate medical and genetics students at the University of Glasgow volunteered to help with illustrations and proofreading; and I am grateful to them all:

- Illustrators: Stefanie Lip, Duncan Henderson, Abigail Little, Hannah Whittingham, Beatrice Tan, Caitlin McNeill, Ananya Das and Greg Marlow.
- Cover design: Dr. Stefanie Lip.
- Proofreading: Nadeem Ahmed, Sonam Ansel, Hnin Aung, Avirup Chowdhury, Ryan Clark, Holli Coleman, Sinead Corbet, Timothy Cryle, Caitlin Doyle, Nora Gonzalez, Manimaalini Kunasegaran, Jun Lau, Prasanna Partha Sarathy, Euan Paterson, Paul Purvis, Elizabeth Twynam-Perkins, Naoise Smyth, and Arthur Woo.

This page is intentionally left blank

Part I

Introduction

This page is intentionally left blank

Chapter 1

Pharmacogenomics and Stratified Medicine

Sandosh Padmanabhan
Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Chapter Outline

1.1 Overview 3
1.2 The Genetic Basis of Drug Response and Adverse Effects 5
1.3 Single-Gene Studies 5
  1.3.1 Predicting Efficacy 5
  1.3.2 Predicting Drug Dosage 10
  1.3.3 Predicting Drug Toxicity 10
1.4 Genome-Wide Association Studies in Pharmacogenomics 12
1.5 The Road to Personalization and Stratification 13
1.6 Pharmacogenetic Biomarkers and Companion Diagnostics 13
  1.6.1 Evaluating Pharmacogenomic Biomarkers/CDx for Clinical Use 17
1.7 Economic Challenges 19
1.8 Conclusions 21
References 22

1.1 OVERVIEW

Drugs are the cornerstone of modern therapeutics. Most of them are currently prescribed empirically based on the physician's assessment of a patient's clinical signs, family history, and probability that a certain drug or therapy will benefit the patient. However, advances in our understanding of the molecular basis of disease are redefining how diseases are classified and present a powerful new dimension for further tailoring treatments to individuals. The principle underpinning drug therapy is that its benefit should outweigh its risks. But in many cases, assessing the balance between risk and benefit may still be a gray area; for example, therapies that provide symptomatic benefits but shorten life may be appropriate in patients with serious and highly symptomatic diseases such as heart failure or cancer.

Large randomized clinical trials (RCTs), meta-analysis, and evidence-based medicine are used to deduce the probability that a drug will work for a specific condition (i.e., provide the benefit of drug therapy). Evidence from RCTs that the drug is appropriately safe and effective is derived statistically from analysis of patient populations participating in these trials; it provides probabilities that are presumed to reflect benefit in the overall population or in specific patient subgroups. Additionally, RCTs define the relationships between dose (or plasma concentration) and benefit–risk ratio, and they establish the population average dose for initiation of therapy. This is the primary requirement for approval of a drug for use. However, it is rare for a drug to be safe or effective for everyone, and individual patients often display striking variability in drug response that requires drug doses that are different from the "average dose." One analysis of drug efficacy from published data showed an 80% response for Cox-2 inhibitors, but the lowest response was 25% for cancer chemotherapy, indicating that in most cancer patients a treatment will only generate severe side effects without treating the disease itself [1,2]. Rates of response to drug treatment are only 30–60% in most common diseases, indicating that the common drug blockbuster strategy ("one drug fits all") may need to be revised according to individual responses [1].

Drug safety varies from drug to drug and from disease to disease, but many drugs have some side effects of clinical

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00001-3
© 2014 Elsevier Inc. All rights reserved.

importance. Approximately 7% of hospital admissions and many post-marketing drug withdrawals are due to adverse drug reactions (ADRs) [3]. Some adverse effects are so common and classically associated with certain drugs that they are identified very early during clinical use. On the other hand, serious adverse effects may be sufficiently uncommon that they escape detection for many years after a drug has been licensed. This is despite intensive efforts by pharmaceutical companies to develop safer drugs and efforts by regulatory agencies to maintain strict safety guidelines [4–6]. Throughout the drug development process, there are two simple realities: people exhibit variable responses to medicines, and unexpected findings are routine [7].

In general, for most commonly used drugs there is no major clinical impact, as many drugs (like nonsteroidal anti-inflammatory drugs or proton-pump inhibitors) work for almost all relevant patients. For other drugs, although responses are variable, there is at present no way to identify patients who are likely to respond well; therefore, a trial-and error approach is followed until a satisfactory response is attained.

Interindividual variation of drug response is caused by a combination of genetic and environmental factors as well as by patient characteristics that can affect the drug's pharmacokinetics and/or pharmacodynamics (Figure 1.1). The effects of some of these variables are obvious (e.g., noncompliance) or addressable (e.g., dose adjustments in patients with impaired kidney function), whereas the mechanisms of many other influences on drug response remain unstudied or unexplained. Data from twin and family studies demonstrate higher interindividual variability than intraindividual variability, indicating the role of shared inherited factors in determining drug response variability [8]. One of the methods for estimating the genetic fraction of a phenotypic variation is to calculate the heritability of the trait. For drug response phenotypes, it is estimated that 20–95% variability in drug disposition and effects may be due to genetic factors [9] (see Chapter 12). These support the notion that information on an individual's predisposition to drug efficacy or toxicity will enable targeted therapy and avoid the drawbacks of the current trial-and-error approach.

Furthermore, rapid advances in genomics and molecular biology are revealing the presence of distinct subpopulations within particular clinical presentations that may be amenable to specific therapies. Classification of patients into these subtypes of disease using diagnostic markers adds another dimension to tailored therapy, where accurate definition of the disease subtype can lead to targeted (in contrast to trial-and-error) treatment.

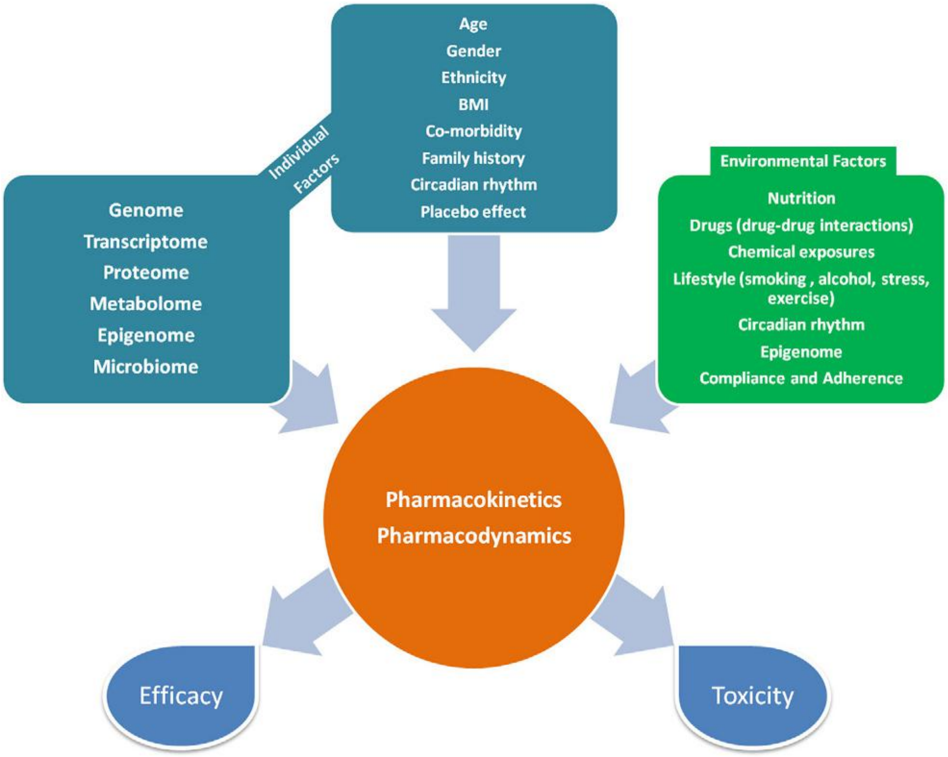

Age
Gender
Ethnicity
BMI
Co-morbidity
Family history
Circadian rhythm
Placebo effect
Individual Factors
Genome
Transcriptome
Proteome
Metabolome
Epigenome
Microbiome
Environmental Factors
Nutrition
Drugs (drug-drug interactions)
Chemical exposures
Lifestyle (smoking , alcohol, stress, exercise)
Circadian rhythm
Epigenome
Compliance and Adherence
Pharmacokinetics
Pharmacodynamics
Efficacy
Toxicity

FIGURE 1.1 Factors involved in interindividual variation in drug response.

BOX 1.1 Historical Aspects of Pharmacogenomics

Recognition of interindividual variability in xenobiotic toxicity stretches back to Pythagoras (c. 580–500 BC) who proscribed fava beans, possibly related to hemolytic syndrome associated with the ingestion of undercooked fava beans. It is now recognized that the underlying cause of this syndrome is a mutation in the gene encoding glucose-6-phosphate dehydrogenase (G6PD), resulting in deficiency of the enzyme. Genetic variation in this gene still has major implications in clinical practice. For example, the recently licensed uricosuric drug rasburicase contains an FDA label warning about G6PD deficiency. In the early 50's phenotype-driven assessment of variation in drug-metabolizing enzyme genes was used to define an individual's N-acetylation capacity as slow or fast (phase II enzyme), whereas debrisoquine hydroxylation was used to define the activity of the phase I cytochrome P450 enzyme (CYP2D6) [10–12].

Subsequent advances in molecular genetics led to the molecular characterization of the defects underlying the debrisoquine hydroxylase or CYP2D6 polymorphism, which included not only polymorphisms but also gene duplication, with individuals having 8–13 copies of the gene [13]. The term "pharmacogenetics" was coined by Friederich Vogel in 1959 [14] to define a new science applying genetics and pharmacology to the study of the influence of inheritance on drug response, but it was only after the human genome project that research in this area accelerated and the timeline is summarized in Figure 1.2.

1.2 THE GENETIC BASIS OF DRUG RESPONSE AND ADVERSE EFFECTS

The term *pharmacogenetics* generally implies the clinical effects of a single or a few gene variants and the associated investigations limited to these loci. It is evident that most genetic variability in drug response is multifactorial, influenced by numerous genes with compensatory or overlapping roles. Consequently, pharmacogenetics has evolved into *pharmacogenomics*, which uses a whole-genome approach, including assessment of gene expression and all pathways to dissect the basis of variability in drug response [14]. The transition from pharmacogenetics to pharmacogenomics was greatly facilitated by the completion of the human genome sequence, the international HapMap project, and advances in high-throughput genotyping, gene expression profiling, next-generation sequencing, and so forth. Pharmacogenomics can be also defined as the combination of pharmacology and genomics at both population and patient levels to advance drug research and development and to manage drug selection and dosing [15].

The first evidence that polymorphisms in drug-metabolizing enzymes have clinical effects was the observation of slow acetylation in relation to isoniazid. Genetic variation in genes involved in drug metabolism may decrease functional activity or expression of the metabolizing enzymes and can give rise to distinct individual phenotypes ranging from poor metabolizer (PM), intermediate metabolizer (IM), extensive metabolizer (EM), and ultrarapid (UM) metabolizer states that have an impact on drug pharmacokinetics and hence drug response (see Chapter 17). Pharmacodynamic pharmacogenomics is the study of altered response to drugs associated with genetic polymorphisms of genes encoding for drug targets and associated modifiers or key proteins (see Chapter 18).

Progress in identifying genetic variation that affects drug response is accelerating, with the FDA including pharmacogenomic information on 38 different genes on relevant drug labels so far (Table 1.1) [16]. However, there are major challenges in translating many of the other associations into clinical use. The influence of some these genetic variations on therapy is summarized below (and in more detail in later chapters).

1.3 SINGLE-GENE STUDIES

1.3.1 Predicting Efficacy

The single-gene approach to targeted therapy can be traced back to as early as 1896, when Sir George Beatson reported clinical improvement in three young women with locally advanced breast cancer following surgical oophorectomy [1]. Landmark studies by Jensen, Lippman, and McGuire, among others, demonstrated that this was due to removal of a specific growth factor, 17β-estradiol, from its cellular target, the estrogen receptor (ER) [17]. Clinicians now routinely use tissue ER content to individualize anti-estrogen therapy for breast cancer patients [18]. The most commonly used adjuvant therapy for the treatment and prevention of estrogen receptor–positive (ER+) breast cancer is Tamoxifen, which results in a 30% reduction in annual breast cancer death compared with placebo [18]. The main anti-estrogenic active metabolite of Tamoxifen is endoxifen, and the key liver enzyme catalyzing the conversion from intermediates to endoxifen is CYP2D6 [19]. Approximately 7% of the population are homozygous for an inactive allele (*4 is the most common variant allele among individuals of European descent, whereas *10 is the most common allele among those of Asian descent); [20] based on the genotype, subjects are categorized as EM, PM, or UM. There is growing clinical evidence of the pharmacogenetic effect of *CYP2D6* variants on tamoxifen treatment outcomes, with PM women showing a significantly shorter time to recurrence [21,22]. Punglia et al. estimated that the 5-year disease-free survival of tamoxifen-treated patients with no mutations (wt/wt) was 83.9%, which is similar to the 84.0% of genotypically unselected patients who were

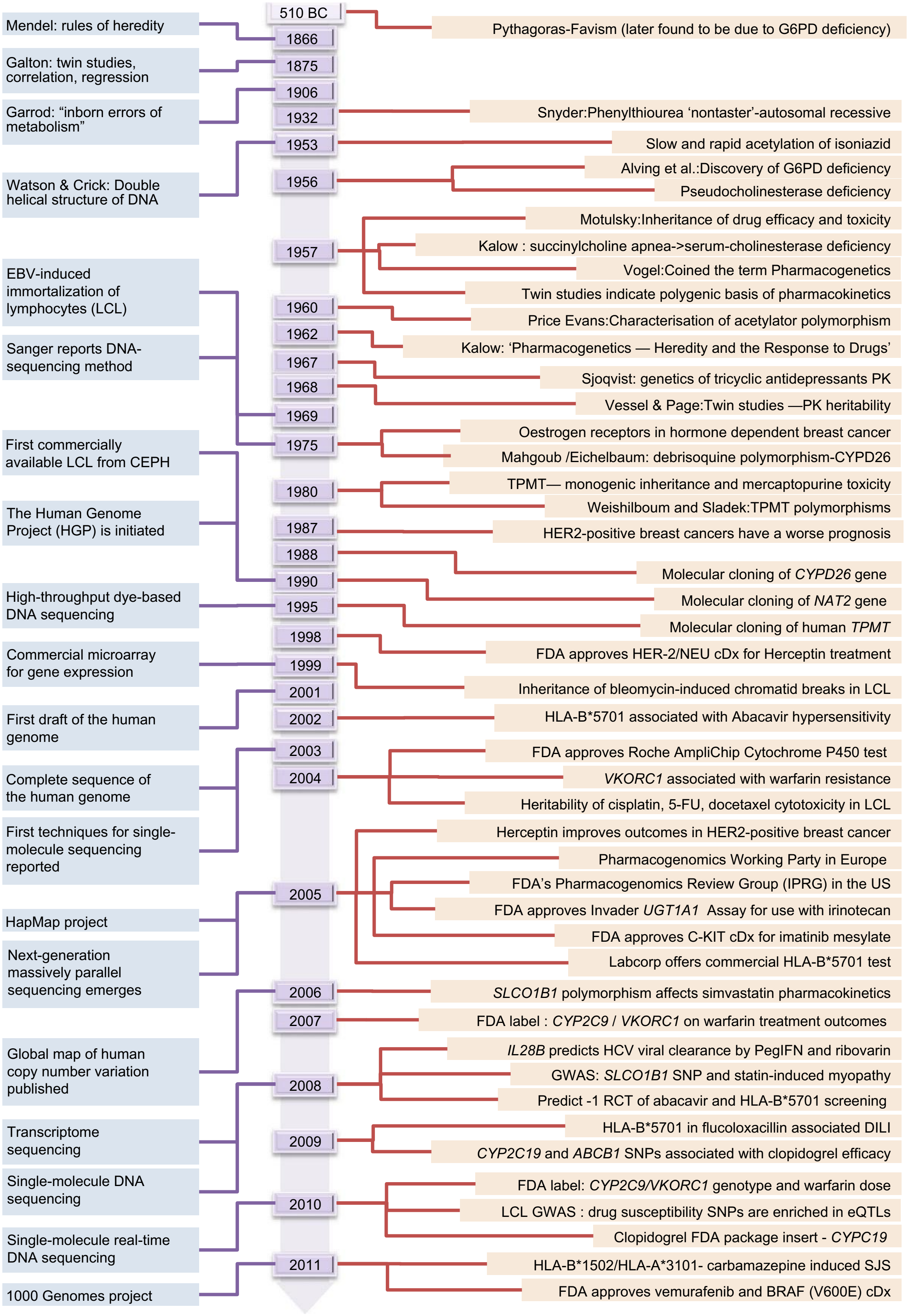

510 BC
Pythagoras-Favism (later found to be due to G6PD deficiency)
Mendel: rules of heredity
1866
Galton: twin studies, correlation, regression
1875
Garrod: "inborn errors of metabolism"
1906
1932
Snyder:Phenylthiourea 'nontaster'-autosomal recessive
Watson & Crick: Double helical structure of DNA
1953
Slow and rapid acetylation of isoniazid
1956
Alving et al.:Discovery of G6PD deficiency
Pseudocholinesterase deficiency
1957
Motulsky:Inheritance of drug efficacy and toxicity
Kalow : succinylcholine apnea->serum-cholinesterase deficiency
Vogel:Coined the term Pharmacogenetics
Twin studies indicate polygenic basis of pharmacokinetics
EBV-induced immortalization of lymphocytes (LCL)
1960
Price Evans:Characterisation of acetylator polymorphism
Sanger reports DNA-sequencing method
1962
Kalow: 'Pharmacogenetics — Heredity and the Response to Drugs'
1967
Sjoqvist: genetics of tricyclic antidepressants PK
1968
Vessel & Page:Twin studies —PK heritability
1969
First commercially available LCL from CEPH
1975
Oestrogen receptors in hormone dependent breast cancer
Mahgoub /Eichelbaum: debrisoquine polymorphism-CYPD26
1980
TPMT— monogenic inheritance and mercaptopurine toxicity
Weishilboum and Sladek:TPMT polymorphisms
The Human Genome Project (HGP) is initiated
1987
HER2-positive breast cancers have a worse prognosis
1988
Molecular cloning of *CYPD26* gene
1990
Molecular cloning of *NAT2* gene
High-throughput dye-based DNA sequencing
1995
Molecular cloning of human *TPMT*
1998
FDA approves HER-2/NEU cDx for Herceptin treatment
Commercial microarray for gene expression
1999
Inheritance of bleomycin-induced chromatid breaks in LCL
2001
First draft of the human genome
2002
HLA-B*5701 associated with Abacavir hypersensitivity
Complete sequence of the human genome
2003
FDA approves Roche AmpliChip Cytochrome P450 test
2004
VKORC1 associated with warfarin resistance
Heritability of cisplatin, 5-FU, docetaxel cytotoxicity in LCL
First techniques for single-molecule sequencing reported
Herceptin improves outcomes in HER2-positive breast cancer
Pharmacogenomics Working Party in Europe
2005
FDA's Pharmacogenomics Review Group (IPRG) in the US
FDA approves Invader *UGT1A1* Assay for use with irinotecan
HapMap project
FDA approves C-KIT cDx for imatinib mesylate
Next-generation massively parallel sequencing emerges
Labcorp offers commercial HLA-B*5701 test
2006
SLCO1B1 polymorphism affects simvastatin pharmacokinetics
2007
FDA label : *CYP2C9* / *VKORC1* on warfarin treatment outcomes
Global map of human copy number variation published
IL28B predicts HCV viral clearance by PegIFN and ribovarin
2008
GWAS: *SLCO1B1* SNP and statin-induced myopathy
Predict -1 RCT of abacavir and HLA-B*5701 screening
Transcriptome sequencing
2009
HLA-B*5701 in flucoloxacillin associated DILI
CYP2C19 and *ABCB1* SNPs associated with clopidogrel efficacy
Single-molecule DNA sequencing
2010
FDA label: *CYP2C9/VKORC1* genotype and warfarin dose
LCL GWAS : drug susceptibility SNPs are enriched in eQTLs
Clopidogrel FDA package insert - *CYPC19*
Single-molecule real-time DNA sequencing
2011
HLA-B*1502/HLA-A*3101- carbamazepine induced SJS
1000 Genomes project
FDA approves vemurafenib and BRAF (V600E) cDx

FIGURE 1.2 Timeline of the major milestones in genetics and pharmacogenetics.

TABLE 1.1 FDA Pharmacogenomic Biomarkers in Drug Labels

Drug	Therapeutic area	Biomarker
Abacavir	Antivirals	HLA-B*5701
Ado-Trastuzumab Emtansine	Oncology	ERBB2 (HER2)
Aripiprazole	Psychiatry	CYP2D6
Arsenic Trioxide	Oncology	PML/RARα
Atomoxetine	Psychiatry	CYP2D6
Atorvastatin	Metabolic and Endocrinology	LDL receptor
Azathioprine	Rheumatology	TPMT
Boceprevir	Antivirals	IL28B
Brentuximab Vedotin	Oncology	CD30
Busulfan	Oncology	Ph Chromosome
Capecitabine	Oncology	DPD
Carbamazepine	Neurology	HLA-B*1502
Carisoprodol	Musculoskeletal	CYP2C19
Carvedilol	Cardiovascular	CYP2D6
Celecoxib	Analgesics	CYP2C9
Cetuximab	Oncology	EGFR, KRAS
Cevimeline	Dermatology and Dental	CYP2D6
Chlordiazepoxide and Amitriptyline	Psychiatry	CYP2D6
Chloroquine	Antiinfectives	G6PD
Cisplatin	Oncology	TPMT
Citalopram	Psychiatry	CYP2C19, CYP2D6
Clobazam	Neurology	CYP2C19
Clomipramine	Psychiatry	CYP2D6
Clopidogrel	Cardiovascular	CYP2C19
Clozapine	Psychiatry	CYP2D6
Codeine	Analgesics	CYP2D6
Crizotinib	Oncology	ALK
Dapsone	Dermatology	G6PD
Dasatinib	Oncology	Ph Chromosome
Denileukin Diftitox	Oncology	CD25
Desipramine	Psychiatry	CYP2D6
Dexlansoprazole	Gastroenterology	CYP2C19, CYP1A2
Dextromethorphan and Quinidine	Neurology	CYP2D6
Diazepam	Psychiatry	CYP2C19
Doxepin	Psychiatry	CYP2D6
Drospirenone and Ethinyl Estradiol	Reproductive	CYP2C19
Eltrombopag	Hematology	Factor V Leiden (FV), SERPINC1
Erlotinib	Oncology	EGFR

(Continued)

TABLE 1.1 Continued

Drug	Therapeutic area	Biomarker
Esomeprazole	Gastroenterology	CYP2C19
Everolimus	Oncology	ERBB2 (HER2)
Exemestane	Oncology	ER &/ PGR
Fluorouracil	Dermatology	DPD
Fluoxetine	Psychiatry	CYP2D6
Fluoxetine and Olanzapine	Psychiatry	CYP2D6
Flurbiprofen	Rheumatology	CYP2C9
Fluvoxamine	Psychiatry	CYP2D6
Fulvestrant	Oncology	ER
Galantamine	Neurology	CYP2D6
Gefitinib	Oncology	EGFR
Iloperidone	Psychiatry	CYP2D6
Imatinib	Oncology	C-Kit, Ph Chromosome, PDGFR, FIP1L1-PDGFRα
Imipramine	Psychiatry	CYP2D6
Indacaterol	Pulmonary	UGT1A1
Irinotecan	Oncology	UGT1A1
Isosorbide and Hydralazine	Cardiovascular	NAT1; NAT2
Ivacaftor	Pulmonary	CFTR (G551D)
Lansoprazole	Gastroenterology	CYP2C19
Lapatinib	Oncology	ERBB2 (HER2)
Lenalidomide	Hematology	Chromosome 5q
Letrozole	Oncology	ER &/ PGR
Maraviroc	Antivirals	CCR5
Mercaptopurine	Oncology	TPMT
Metoprolol	Cardiovascular	CYP2D6
Modafinil	Psychiatry	CYP2D6
Mycophenolic Acid	Transplantation	HGPRT
Nefazodone	Psychiatry	CYP2D6
Nilotinib	Oncology	Ph Chromosome, UGT1A1
Nortriptyline	Psychiatry	CYP2D6
Omeprazole	Gastroenterology	CYP2C19
Panitumumab	Oncology	EGFR, KRAS
Pantoprazole	Gastroenterology	CYP2C19
Paroxetine	Psychiatry	CYP2D6
Peginterferon alfa-2b	Antivirals	IL28B
Perphenazine	Psychiatry	CYP2D6
Pertuzumab	Oncology	ERBB2 (HER2)

TABLE 1.1 Continued

Drug	Therapeutic area	Biomarker
Phenytoin	Neurology	HLA-B*1502
Pimozide	Psychiatry	CYP2D6
Prasugrel	Cardiovascular	CYP2C19
Pravastatin	Metabolic and Endocrinology	APOE2
Propafenone	Cardiovascular	CYP2D6
Propranolol	Cardiovascular	CYP2D6
Protriptyline	Psychiatry	CYP2D6
Quinidine	Antiarrhythmics	CYP2D6
Rabeprazole	Gastroenterology	CYP2C19
Rasburicase	Oncology	G6PD
Rifampin, Isoniazid, and Pyrazinamide	Antiinfectives	NAT1; NAT2
Risperidone	Psychiatry	CYP2D6
Sodium Phenylacetate and Sodium Benzoate	Gastroenterology	UCD (NAGS; CPS; ASS; OTC; ASL; ARG)
Sodium Phenylbutyrate	Gastroenterology	UCD (NAGS; CPS; ASS; OTC; ASL; ARG)
Tamoxifen	Oncology	ER, Factor V Leiden (FV), Prothrombin mutations (F2)
Telaprevir	Antivirals	IL28B
Terbinafine	Antifungals	CYP2D6
Tetrabenazine	Neurology	CYP2D6
Thioguanine	Oncology	TPMT
Thioridazine	Psychiatry	CYP2D6
Ticagrelor	Cardiovascular	CYP2C19
Tolterodine	Reproductive and Urologic	CYP2D6
Tositumomab	Oncology	CD20 antigen
Tramadol and Acetaminophen	Analgesics	CYP2D6
Trastuzumab	Oncology	ERBB2 (HER2)
Tretinoin	Oncology	PML/RARα
Trimipramine	Psychiatry	CYP2D6
Valproic Acid	Psychiatry	UCD (NAGS; CPS; ASS; OTC; ASL; ARG)
Vemurafenib	Oncology	BRAF
Venlafaxine	Psychiatry	CYP2D6
Voriconazole	Antifungals	CYP2C19
Warfarin	Hematology	VKORC1, CYP2C9

treated with aromatase inhibitors [23]. Chapter 17 discusses these issues in detail.

Proton-pump inhibitors (PPIs) are among the most extensively used drugs in the world. In EMs, CYP2C19 is responsible for >80% of the metabolism of omeprazole, lansoprazole, and pantoprazole [24], with CYP3A4 undertaking most of the remaining metabolism [25,26]. The metabolites produced are inactive. Eight variants (*2–*8, *17) in the *CYP2C19* gene have now been associated with reduced enzyme activity *in vivo*, largely because of the production of inactive enzyme proteins [27,28]. Standard recommended doses of any of these PPIs might not provide sufficient acid suppression in homozygous EMs, suggesting that an increase in dose might be required in this group. Indeed, one study showed that increasing the lansoprazole 30 mg dose from once daily to four times daily in homozygous EMs leads to an increase in mean 24 h intragastric pH from 4.5 to 7.0 [29]. (see Chapter 17)

1.3.2 Predicting Drug Dosage

Warfarin pharmacogenomics offers the best example of genetic variation influencing drug dosage that is validated across different racial/ethnic/geographic groups (see Chapter 24). The effect on warfarin dose requirements of variants of cytochrome P450 2C9 (CYP2C9—the principal enzyme in warfarin metabolism) and vitamin K epoxide reductase complex 1 (VKORC1—the target protein inhibited by warfarin to produce therapeutic anticoagulation) has been extensively studied, with 2–10% and 10–25% of the variance in dose independently accounted for by the *CYP2C9* and *VKORC1* genotype, respectively [30–32]. A pharmacogenetic warfarin dose refinement algorithm based on clinical information, INR, and genetic factors was shown to explain at least 69.1% of therapeutic dose variability after about one week of therapy [33]. Subsequently, an algorithm produced by the International Warfarin Pharmacogenomics Consortium showed that genotype-guided therapy exhibits the greatest benefits in patients, ultimately requiring 3 mg/d or less (33.9% of cohort) or 7 mg/d or more (12.4% of cohort). However, the first prospective, randomized controlled trial to compare genotype-guided dosing to "standard" empirical dosing has produced a negative result [34]. This has led to caution about warfarin genetic testing:

> *"Testing could result in net harm to patients. For example, a result of "wild type" from the genetics laboratory may impart an unjustified sense of security to the clinician, leading to inappropriately high initial warfarin doses, an unsafe delay in early follow-up INR measurements for a patient new to warfarin, or an inappropriately long interval between INR checks for warfarin-experienced patients. Similarly, a genotype that suggests that a patient needs low doses may lead to initial underdosing with resultant subtherapeutic INR, exposing the patient to an avoidable risk of thrombosis. Any such unintended adverse effects from genetic testing would need to be subtracted from any hypothetical benefits in a proper cost-effectiveness calculation" [35].*

1.3.3 Predicting Drug Toxicity

The association of prolonged muscle paralysis and apnea by the muscle relaxant succinylcholine with atypical butyrylcholinesterase (pseudocholinesterase) is a well-known highly penetrant monogenic pharmacogenomic trait (see Chapters 18 and 35). Four percent of Caucasians have at least one abnormal *BCHE* allele that results in an enzyme with either altered affinity (atypical (A), fluoride (F), and some silent (S) variants) or decreased quantity (e.g., K, J, and some S variants). Only 1 in 3500, or ~0.03%, have genotypes (AA, AK/AK, A/AK, SS, S/AK, and AS) that result in BChE activity that is low enough to prolong apnea with a muscle relaxant. Screening for the five most common mutations (fluoride-1, fluoride-2, K variant, dibucaine, and silent-1) identifies 85% of individuals with a genetic basis for prolonged apnea with neuromuscular blockers [36].

Isoniazid is an effective treatment for tuberculosis, but is limited by peripheral neuropathy. It was observed that individuals at greatest risk of peripheral neuropathy excrete more isoniazid unchanged in the urine and a reduced amount of acetylisoniazid [11], suggesting a putative role for a drug-metabolizing enzyme. The enzyme responsible for isoniazid acetylation is N-acetyltransferase type 2 (NAT2), and variation in the gene affects acetylation capacity, resulting in a bimodal distribution of isoniazid plasma concentrations [37] attributable to a single-gene effect. NAT2 also acetylates a number of other drugs, including hydralazine, procainamide, dapsone, and some sulfonamide derivatives. Slow acetylator status is associated with a higher risk of peripheral neuropathy with isoniazid and a lupus-like reaction with hydralazine [38] (see Chapters 17 and 36).

Carbamazepine is a widely used drug for the treatment of epilepsy, bipolar disorder, trigeminal neuralgia, and chronic pain. A rare serious adverse effect of carbamazepine is Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN), which is associated with 30% mortality [39,40]. A case–control study conducted on Han Chinese residing in Taiwan first identified HLA-B*15:02 as the genetic marker for CBZ-induced SJS/TEN [41]. HLA-B*15:02 is largely absent in individuals not of Asian origin (e.g., Caucasians, African Americans, Hispanics, and Native Americans). Although no published data have confirmed the HLA-B*15:02–SJS/TEN association, the higher prevalence of this variant in Asian populations [42] resulted in the Food and Drug Administration's decision to recommend testing in patients of Asian origin. (see Chapters 21).

Thiopurine methyltransferase (TPMT) is a cytosolic drug–metabolizing enzyme that probably makes the strongest case of all for prospective pharmacogenetic testing

(Figure 1.3). TPMT catalyzes the S-methylation of 6-mercaptopurine, azathioprine, and thioguanine to inactivate the thiopurine drugs, which are used for the treatment of leukemia and autoimmune diseases. More than 20 variant alleles of the TPMT gene have been identified, among which TPMT*2, TPMT*3A, and TPMT*3C are defective alleles that produce poor enzymatic activities. Inherited TPMT deficiency predisposes to myelosuppression with both azathioprine and its initial product, 6-mercaptopurine (6-MP) [43]. In Caucasians, a trimodal distribution exists, with 0.3–0.6% having low or undetectable activity (homozygotes for the variant allele, e.g., *3A/*3A), 10% having intermediate activity (heterozygotes, e.g., TPMT*1/*3A), and the remaining 90% having high (normal) activity (homozygotes for wild-type alleles, e.g., *1/*1) [44,45]. TPMT*3A is the most common variant in Caucasians [46], whereas *3C is the most frequent mutation in those of African [47] or East or Southeast Asian origin [48,49]. One study showed that screening for TPMT*2 and *3A – *3D plus sequencing of the open reading frame predicts low, intermediate, or normal/high TPMT activity phenotype in 98% of cases [45] (see Chapter 17).

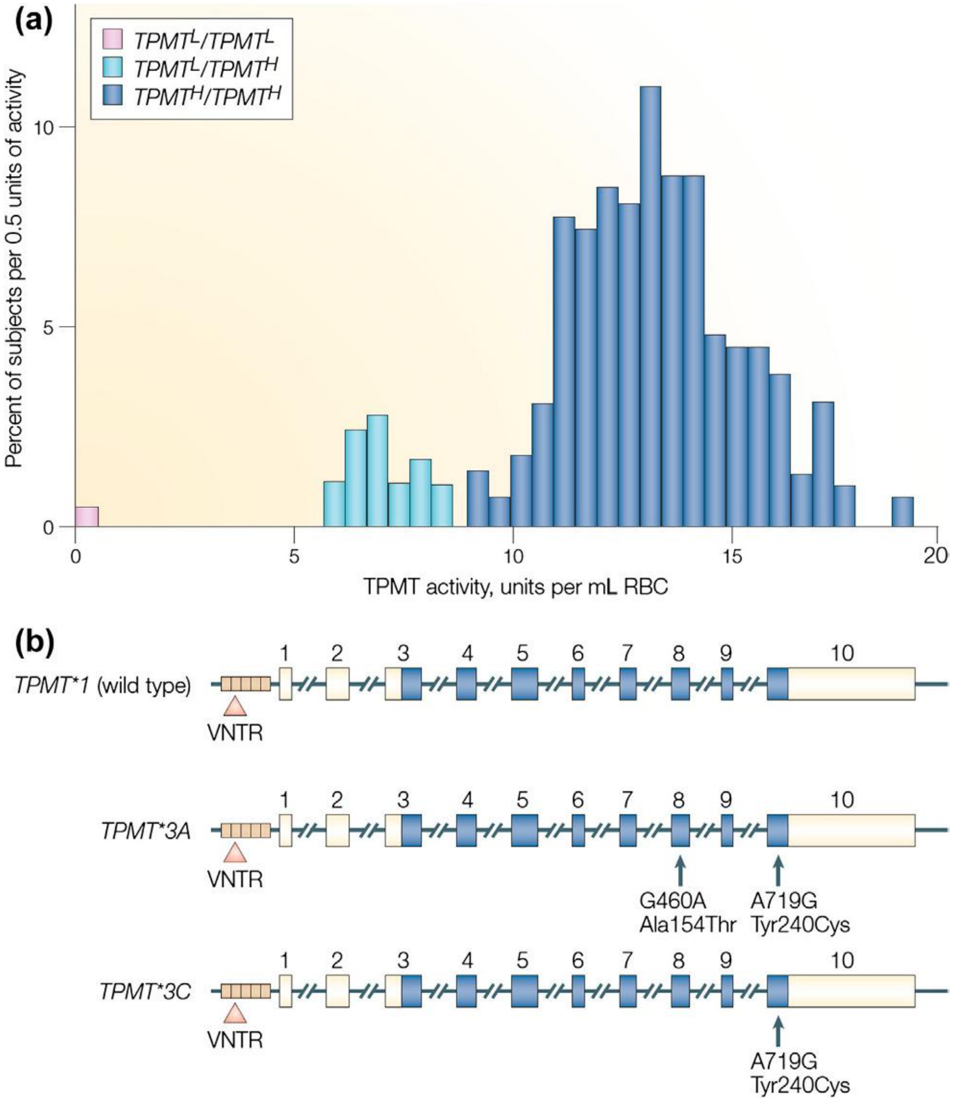


FIGURE 1.3 Classic pharmacogenetic traits: thiopurine S-methyltransferase polymorphism. (a) Activity of the drug-metabolizing enzyme thiopurine S-methyltransferase (TPMT) in red blood cells from 298 randomly selected Caucasian blood donors. Presumed genotypes for the TPMT genetic polymorphism are also indicated. $TPMT^L$ and $TPMT^H$ are designations for alleles resulting in "low" and "high" activity, respectively. These allele designations were used before the molecular basis for the polymorphism was understood. (b) TPMT alleles. TPMT*1 is the most common allele (wild-type) generally; TPMT*3A, with two nonsynonymous-coding SNPs, is the most common variant allele in Caucasian subjects; and TPMT*3C is the most common variant allele in East Asian subjects [46]. Rectangles represent exons, with blue areas representing the open reading frame. Arrows indicate two SNPs, as well as a polymorphic variable number of tandem repeats (VNTR) in the promoter. *Source: (a) Modified with permission from* [44], *© University of Chicago Press (1980). Reprinted with permission from* National Review of Drug Discovery [15].

Uridine diphosphate glucuronosyltransferase 1A1 (UGT1A1) belongs to the uridine diphosphate glucuronosyltransferase superfamily, and it is the first pharmacogenetic test to achieve FDA approval for use in conjunction with a specific drug (irinotecan). The activity of UGT1A1 varies widely, with an in vitro study demonstrating a 17-fold variation in SN-38 glucuronidation [50]. High levels of SN-38 lead to severe side effects, including severe myelosuppression in 15–20% and severe delayed-type diarrhea in 20–25% of patients receiving irinotecan. UGT1A1*28 is the variant most frequently implicated in defective SN-38 (the active metabolite of irinotecan) glucuronidation, and it occurs commonly; the homozygous genotype is found in 5–15% of Europeans, 10–25% of Africans and South Asians, and 1–5% of Southeast Asians and Pacific Islanders [51]. The FDA has recommended that patients be genotyped for the UGT1A1*28 polymorphism and that the dose be adjusted accordingly before irinotecan treatment (see Chapter 17).

1.4 GENOME-WIDE ASSOCIATION STUDIES IN PHARMACOGENOMICS

Candidate gene approaches have resulted primarily in the identification of genetic variants in drug-metabolizing genes with large effects on toxicity or response. Genome-wide association studies (GWAS) have identified a number of common single-nucleotide polymorphisms (SNPs) associated with drug efficacy and safety, with the majority of the signals in unanticipated genes or in loci of unknown functional relevance. When considering all of the GWAS cataloged in the U.S. National Human Genome Research Institute, pharmacogenomics studies are sevenfold more likely to achieve odds ratios (ORs) >3.0 compared with common disease GWAS [52]. Two major reasons have been proposed for the large effect of single-nucleotide polymorphisms (SNPs) identified in pharmacogenomic studies [52]. First, until recently humans had not been exposed to synthetic drugs—modern drugs with the longest histories have been available for only a little over 100 years—and so there has been little negative evolutionary pressure for drug use, allowing polymorphisms to become relatively common in human populations. Second, there might be a large interplay between the gene and the drug effect. For example, carbamazepine-induced severe skin-adverse reaction [41] manifests only if an individual takes the drug. The size of the subject population required to obtain significantly different allele frequencies in individuals with the phenotype versus those without might be expected to be relatively small [52]. Table 1.2 summarizes the major findings from pharmacogenomic GWAS, highlighting some of the novel pathways identified using this approach.

TABLE 1.2 Associations Between SNPs and Drug Toxicity/Efficacy From Genome Wide Association Studies

Drug	Response/toxicity	Lowest p-value	Gene	Reference
Adverse Reaction				
Ximelagatran	Drug-induced liver injury	6.0×10^{-6}	*HLA-DRB1*0701;DQA1*0201*	[53]
Flucloxacillin	Drug-induced liver injury	8.7×10^{-33}	*HLA-B*5701*	[54]
Lumiracoxib	Drug-induced liver injury	2.8×10^{-10}	*HLA-DRB1*1501; DQB1*0602*	[55]
Amoxicillin-clavulanate	Drug-induced liver injury	4.8×10^{-14}	*HLA-DRB1*1501; DQB1*0602;HLA-A*0201*	[56]
Carbamazepine	Stevens-Johnson Syndrome	1.2×10^{-13}	*HLA-A*3101;HLA-B*1502*	[57,58]
Simvastatin	Myotoxicity	4.0×10^{-9}	*SLCO1B1*	[59]
Iloperidone	QT prolongation	2.8×10^{-6}	*CERKL*	[60]
Pamidronate, zoledronic acid	Osteonecrosis of the jaw	6.2×10^{-6}	*CYP2C8*	[61]
Interferon a2b/ribavirin	Hemolytic anemia	4.6×10^{-52}	*ITPA*	[62]
Dabigatran	Bleeding	1.2×10^{-8}	*CES1*	[63]
Olanzapine, perphenazine, quetiapine, risperidone, and ziprasidone	QT prolongation	1.5×10^{-7}	*SLC22A23*	[64]
PEG-IFN/RBV	Thrombocytopenia	8.2×10^{-9}	*ITPA/DDRGK1*	[65]
Cisplatin	Cytotoxicity	1.6×10^{-7}		[66]
Nevirapine	Rash	1.6×10^{-4}	*CCHCR1*	[67]

TABLE 1.2 Continued

Drug	Response/toxicity	Lowest *p*-value	Gene	Reference
Aromatase inhibitors	Musculoskeletal adverse events	6.6×10^{-7}	*TCL1A*	[68]
Carboplatin	Carboplatin sensitivity (cell model), progression-free survival (patients)	9.8×10^{-6}	*ALDH2*	[69]
Antipsychotics	Extrapyramidal symptoms	8.9×10^{-6}		[70]
Citalopram	General side effect burden, overall tolerability, sexual side effects, dizziness and vision/hearing side effects	3.2×10^{-8}	*EMID2*	[71]
Epirubicin	Leukopenia	1.6×10^{-7}	*MCPH1*	[72]
Allopurinol	Stevens-Johnson syndrome, toxic epidermal necrolysis	2.4×10^{-8}	*BAT1, HCP5, and MICC*	[73]
Aspirin	Aspirin-intolerant asthma	6.0×10^{-5}	*CEP68*	[74]
Drug Response				
Gemcitabine plus either bevacizumab or placebo	Overall survival	9.5×10^{-7}	*(IL)17F*	[75]
Citalopram	Clinical scoring for depression response	5×10^{-7}		[76]
Angiotensin-converting enzyme (ACE) inhibitors	Blood pressure response	3.0×10^{-25}	*ABO*	[77]
Candesartan	Blood pressure change	8.9×10^{-7}	*FUT4*	[78]
Hydrochlorothiazide	Blood pressure change	3.3×10^{-8}	*PRKCA; GNAS-EDN3*	[79]
Glucocorticoids	Response to glucocorticoid therapy in asthma	7×10^{-4}	*GLCC11; T*	[80,81]
Warfarin (Coumadin)	Maintenance dose	$<10^{-10}$	*VKORC1; CYP2C9; CYP4F2*	[30,32]
Clopidogrel (Plavix)	Anti-platelet responsiveness	1.5×10^{-13}	*CYP2C19*	[82]
Interferon a (Pegasys)	Sustained hepatitis C genotype 1 viral response	1.06×10^{-25}	*IL28B*	[83]
Metformin	Glycemic response	2.9×10^{-9}	*ATM*	[84]
Methotrexate	Pharmacokinetics	1.7×10^{-10}	*SLCO1B1*	[85]

1.5 THE ROAD TO PERSONALIZATION AND STRATIFICATION

Accelerated developments in genomics and other high-throughput technologies, cellular programming, and stem cells are having a major impact not only in drug development but also in the practice of medicine. The current clinical drug development process is fraught with uncertainty, with up to 90% of human trials on new drugs failing to achieve registration [86]. It is evident that identifying patients who are probable nonresponders to the standard dosing regimen and/or patients with an increased risk of adverse drug reactions will have a positive impact on the benefit–risk ratio of a drug and consequently lead to effective drug development.

1.6 PHARMACOGENETIC BIOMARKERS AND COMPANION DIAGNOSTICS

In order to translate pharmacogenomic discoveries to the clinical setting, a biomarker or a diagnostic test is essential.

BOX 1.2 Therapeutic Continuum

Empirical Medicine

Current clinical practice can be termed “all comer,” or “empirical,” medicine, where the basis of treatment is derived not from the scientific method but rather from practical experience or observation, such as clinical trials or meta-analysis. Empirical medicine relies on the assumption that population-based response data from clinical trials can be applied to the individual patient, and in many cases this is reasonable where the drug works for most patients (e.g., NSAIDs). However, when drugs are effective in only a subset of patients (e.g., antidepressants) and no method is available to identify these patients, treatment is administered to everyone with the diagnosis.

Individualized or Personalized Medicine

At the other end of the continuum is individualized or personalized medicine, where therapy is matched specifically to an individual patient and may not work in other patients. In contrast to empirical medicine, this generally promises to result in both higher-quality treatment for individual patients and lower healthcare costs since patients will be offered only therapies that are effective for them and nonsafe or noneffective treatments will be avoided.

Most advances in individualized medicine come from oncology. For example, the cancer vaccine Oncophage, which is currently in development, utilizes a unique “signature” obtained from tumor cells taken from a patient during surgery [87] for the production of a vaccine that is administered to the patient post-surgery. This vaccine, which is only suitable for the patient from whom it is derived, stimulates an immune response that attacks tumor cells remaining after surgery in the body.

Stem-cell–based therapies could also potentially increase the role of individualized medicine in the future. Cellular reprogramming is technically laborious and can take several weeks, which limits the application of this technology [88]. The recent demonstration that various solid tumors arise and are maintained by cancer stem cells (CSC) advocates the development of treatments targeting CSCs in order to enhance cancer therapy effectiveness and reduce tumor relapse [89]. Despite the promise of induced pluripotent stem (iPS) cells, several limitations currently exist — reactivation of the oncogene c-Myc, immune rejection, inefficiency of reprogramming, epigenetic memory, and so forth [89]. Less extreme examples of patient matching to a medicine include therapies for which some form of marker is available to indicate that a given patient is likely to show a response to the therapy.

Stratified Medicine

Current pharmacological treatment, or empirical treatment, is based on a stepped or adaptive model of care where tailoring of treatment depends on the patient’s actual response to previously offered treatment, with a sequence of interventions (which may differ in intensity, duration, cost, or complexity) being offered to those who have not responded sufficiently [90,91]. Stratified medicine adds a step to traditional clinical practice, and instead of proceeding to treat all patients, it uses baseline information about a patient’s likely response to formulate treatment decisions. In stratified medicine, treatment (including pharmacological and nonpharmacological interventions) is targeted to specific subgroups of patients who share certain biological or risk characteristics, thereby generating more favorable benefit–risk profiles for the therapeutics than with a traditional empirical “all-comers” or empirical approach [92,93].

Stratified medicine is already practiced in several contexts. Examples in common clinical practice include the decision to give lipid-lowering therapy to people above a certain threshold of cardiovascular risk estimated from a prognostic model; the use of bisphosphonates for women over the age of 50 considered to have an increased risk of vertebral fractures; the targeting of primary care management of back pain; antibiotics matched to specific resistant infections; HIV therapies tailored to the viral variant present; a positive HER-2 status in breast cancer; and the use of trastuzumab or imatinib in patients with chronic myeloid leukemia targeted to those with the BCR-ABL mutation; and gefitinib used to treat pulmonary adenocarcinoma in patients with epidermal growth factor receptor mutations [94,95]. Individuals of African ancestry tend to have a salt-sensitive type of hypertension that is responsive to diuretics or calcium channel blockers, while Europeans are more responsive to ACE inhibitors or angiotensin receptor blockers; initial management of hypertension uses a stratified approach based on ethnicity [96,97]. Stratified medicine can therefore increase medical value both for treated patients and for untreated patients, who can more rapidly seek alternative therapies that may be more effective.

The transition from empirical medicine to stratified medicine occurs when biomarkers are identified that can classify patients based on therapeutic efficacy or adverse effects. This is the stratification step that occurs after a disease diagnosis is made and treatment options are considered. While stratified medicine lies in between empirical and individualized medicine in the therapeutic continuum (see Figure 1.4) [92], sometimes, the stratification biomarker identifying differential response can lead to the identification of a new subtype of the index disease. This can result in the stratified medicine patient subpopulation emerging as a clinically distinct diagnosis treated empirically with the stratification biomarker now used as a diagnostic test. For example, in breast cancer expression of the receptor tyrosine kinase HER2 characterizes the cancer type and the suitability for receiving trastuzumab (Herceptin, Genentech) as an adjuvant therapy [95].

Pharmacogenomics and stratified medicine are part of the continuum of drug development and clinical implementation toward reduced dependence on “trial-and-error medicine” as well as improved efficiency and effectiveness of healthcare delivery [92,98]. Stratification represents a more targeted approach to therapy, offering the potential of greater efficacy of treatments and minimization of their side effects. For example, the 2–9% of patients who test positive for HLA-B*57:01 will not be prescribed Abacavir, leading to decreases in the requirement to treat illness arising from adverse drug reactions, which could be of great benefit to healthcare systems in limiting costs. Another example is the 55% efficacy of crizotinib in patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) that is anaplastic lymphoma kinase– (ALK-) positive, compared to the approximately 5% overall response.

Not surprisingly, pharmacogenomic discoveries are often followed by discussions about clinical implementation. The cost-effectiveness of stratification will depend on several factors, including the cost of testing the patient population, the proportion of "responders," and the (saved) costs of treating nonresponders, which will be greater as the health consequences of any adverse reaction grow. For a particular disease, a stratified medicine strategy can emerge only when certain conditions exist, as shown in Figure 1.5.

More importantly, it is essential that two conditions be met: the presence of multiple treatment options for the disease with heterogenous responses and a logistically and medically acceptable biomarker that can stratify different treatment options [92]. If no biomarker exists, stratified medicine will not emerge because a therapy and a patient subpopulation cannot be linked. However, identification of markers associated with drug response does not always equate to clinically useful predictors. The translation and acceptance of novel genetic and genomic biomarkers in clinical practice must be dependent upon the evidence associated with individual markers to be useful for a particular purpose.

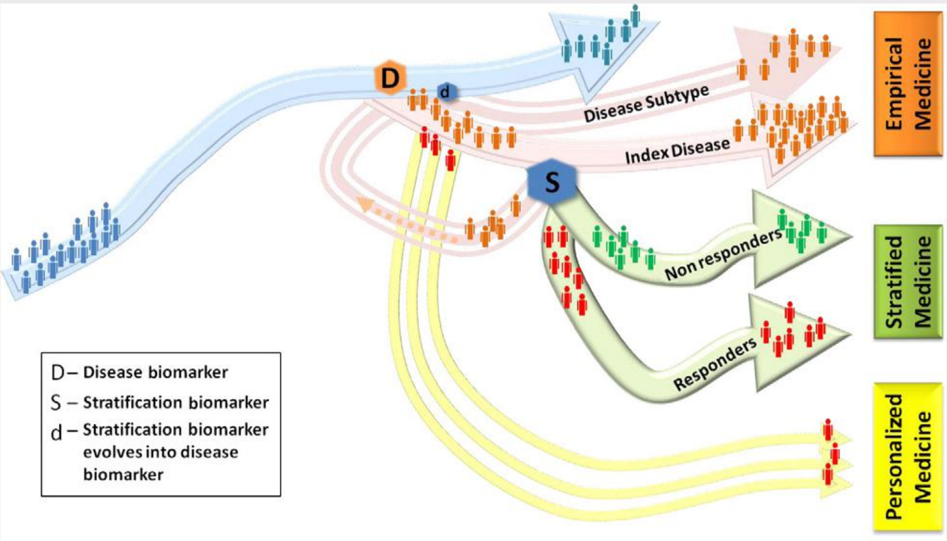

D
d
S
Disease Subtype
Index Disease
Non responders
Responders
Empirical Medicine
Stratified Medicine
Personalized Medicine
D – Disease biomarker
S – Stratification biomarker
d – Stratification biomarker evolves into disease biomarker

FIGURE 1.4 Therapeutic continuum: empirical, stratified, and personalized medicines.

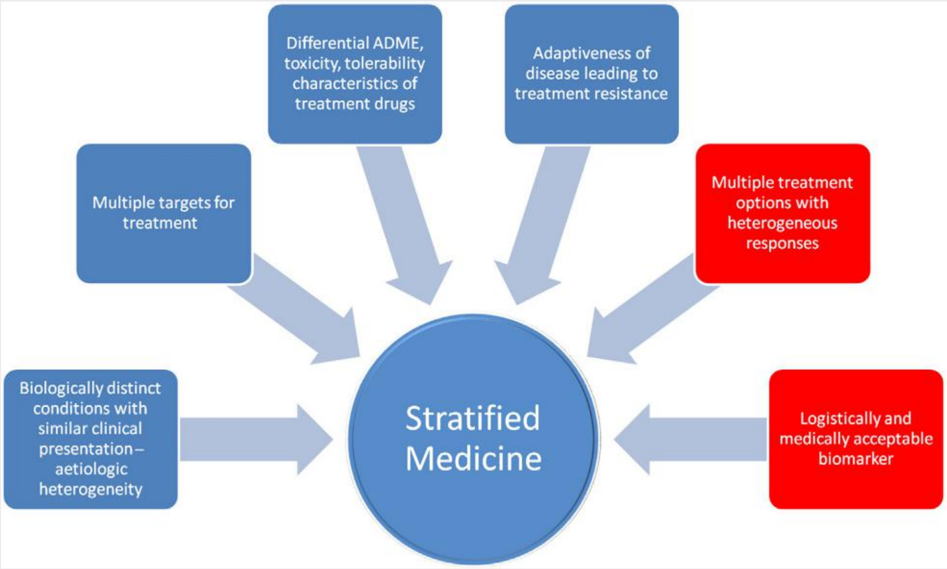

Differential ADME, toxicity, tolerability characteristics of treatment drugs
Adaptiveness of disease leading to treatment resistance
Multiple targets for treatment
Multiple treatment options with heterogeneous responses
Biologically distinct conditions with similar clinical presentation– aetiologic heterogeneity
Stratified Medicine
Logistically and medically acceptable biomarker

FIGURE 1.5 Necessary conditions for stratified medicine. *Source: Adapted from* [92].

A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacological responses to a therapeutic intervention [99]. A pharmacogenomic biomarker is any molecular "barcode" obtained by analysis of mRNA, DNA, protein, or circulating cells that may be used to stratify patients for treatment benefit within clinical trials, to prognosticate patient outcome, or to predict and/or monitor response to therapy.

The result of the pharmacogenomic biomarker assay can determine (1) whether a patient is suitable for treatment with a particular drug (by stratification markers), (2) the most effective dose for the patient (by efficacy markers), (3) the underlying susceptibility of a patient to a particular side effect or group of side effects (by toxicity markers), or (4) the course and effectiveness end point of a therapy (by surrogate end point markers) [100–102]. The goal of these tests is to change the benefit–risk ratio to a degree that would justify the cost and inconvenience of the test (Figure 1.6).

While the FDA website lists ~122 different biomarkers such as gene variants, functional deficiencies, expression changes, and chromosomal abnormalities (see Table 1.1), pharmacogenomic information with action-oriented instructions for physicians and patients is presently limited to approximately 25–30 therapeutic agents, depending on healthcare system and country.

Pharmaceutical companies are increasingly looking to develop a drug and diagnostic test (companion diagnostic, or CDx) simultaneously (drug-diagnostic/co-development) to stratify the patient population for treatment [103]. There are multiple technology platforms that can aid the discovery of companion diagnostics; these are summarized in Table 1.3. Until recently, pharmaceutical companies were focused on "blockbusters" and were unsupportive of "companion diagnostics," thinking that these would divide patient populations into smaller groups and contract their markets. Moreover, early knowledge of relevant predictive biomarkers has historically been the exception rather than the rule in drug development, and the potential for stratified medicine so far has not been fully realized [104,105]. Following the success of two targeted anticancer drugs—trastuzumab (Herceptin; Roche/Genentech) and imatinib (Gleevec;

TABLE 1.3 Biomarker Discovery Technology Platforms and Potential for Companion Diagnostic Utility

Technology platform	Examples	Characteristics for assay development	Clinical utility
Mass spectrometry	Metabolomics	• Low throughput • Expensive • No automation	• Low throughput • Costly
Protein recognition assays (immunoassays)	Immunoassays, aptamers	• High sensitivity • Rapid • High throughput	Limited to certain therapeutic areas
Imaging	Microscopy, computed tomography (CT), Magnetic resonance tomography (MRT), Near-infrared spectroscopy (NIRS), photo-electron tomography (PET)		Useful as prognostic biomarker of drug impact on tumor size as opposed to a metric of tumor shrinkage
Polymerase chain reaction (PCR) as an assay	Allele-specific PCR, qPCR, digital PCR (dPCR)	• Requires ≥ 1 μg genomic DNA • Long assay development time	• Standard for many assays • Useful as setup for sequencing • Low cost • High throughput
Next-generation (NGS) sequencing	Whole genome, targeted, hybridization selection, amplicon sequencing	• Require nanogram amounts of DNA • Faster assay development time • Low cost/base • High sensitivity • High throughput • Flexible to quickly incorporate changes	• High throughput • Multiplex capability • Infinite number of regions to target in single run • Single run can provide data useful in heterogeneous and complex oncology diseases
Anatomic histopathology	Immunohistochemistry (IHC), in situ hybridization (ISH), quantum dots, fluorescence, resonance, or radioactive labeling; multiplex IHC	• Allows specific target localization provided antibodies have high specificity • Ideal way to detect protein pathway activation • Long assay development time	• Limited multiplex capabilities

Source: Adapted from Love D et al. [107].

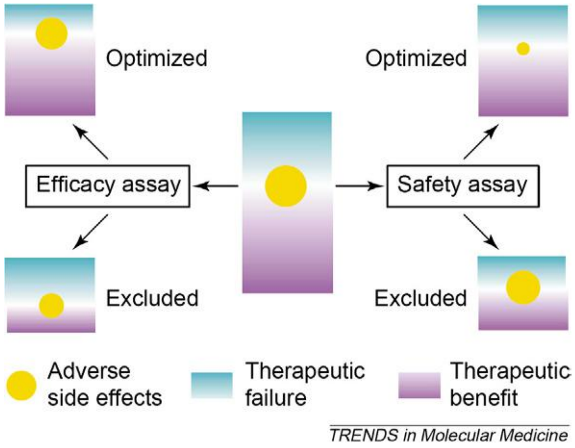


FIGURE 1.6 Clinical application of pharmacogenetic tests. In the general population (center) some individuals taking a particular drug will derive therapeutic benefit (purple) and some will not (blue). Also, some individuals will have a characteristic adverse side effect (yellow). Pharmacogenetic assays will determine whether a patient is more or less well suited to the particular drug based on results from a genotyping assay. In some cases, there might be both efficacy- and safety-based assays. *Source: Reprinted with permission from Elsevier: Trends in Molecular Medicine* [1].

Novartis)—which combined a drug with a companion diagnostic to select or exclude patients for treatment, there is now a growing list of FDA-approved therapies and their companion diagnostics (Table 1.4). Additionally, CDxs can make drug development more efficient and effective because they can reduce costs through preselected (smaller) patient populations, improve the chances of approval, and significantly increase market uptake [104,105]. In July 2011, the FDA released draft guidelines that require CDx tests to be approved simultaneously with their accompanying therapies if necessary for the safe and effective use of a medication. Developing CDx tests in conjunction with therapeutics allows the design of better clinical trials and eventually better implementation of pharmacogenomics testing [106].

1.6.1 Evaluating Pharmacogenomic Biomarkers/CDx for Clinical Use

Recently Love et al. [107] proposed seven stages of CDx co-development, providing an overall depiction of the key questions to address at each phase of CDx development.

Defining an unmet need. What is the clinical problem to be solved? How the test and therapeutic contribute to the solution should be articulated in a target product profile (TPP) for both the test and the therapeutic. Moreover, it is of crucial importance to establish criteria by which success for both will be measured in terms of sufficiently fulfilling the TPP.

Biomarker discovery. What is available to solve the problem? During this stage one must characterize the information provided by a biomarker (or multiple biomarkers), determine the appropriate platform on which to identify the analyte, recognizing that it could be *in vivo* imaging or an in vitro analyte and selecting the appropriate body "compartment" from which to extract the information (e.g., sample type).

Technical assay validation: Can you reliably measure the biomarker(s)? The key deliverable at this stage is demonstrating that the test and any associated instrumentation are capable of reproducibly measuring the biomarker(s) in relevant sample type(s).

Clinical validation: Is the test informative and relevant in the clinical condition/disease? At this stage, the test is evaluated to determine if its performance displays an appropriate degree of reproducibility (precision and accuracy), sensitivity, specificity, and linearity over the range of possible test results in light of the clinical condition for the drug's intended use in the TPP.

Clinical utility. Does the test provide actionable information regarding the condition/disease? Data obtained during this stage demonstrate how the positive and negative predictive value provided by the test enables improved treatment decision making, typically by further refining the characterization of the patient's condition/disease beyond what is possible with the current standard of care.

Medical utility. Is the information used to make a treatment decision that improves an outcome? At this stage, the test typically enters routine usage by medical professionals, who use the information obtained from it to make treatment decisions. The test is often incorporated into treatment guidelines and demonstrates an impact on treatment outcomes.

Commercial adoption. Is the test and associated treatment established as a standard of care and appropriately reimbursed? In this final stage, the test-and-treatment approach is fully integrated into routine patient care and often data exist suggesting or demonstrating a health/economic benefit from the use of this approach.

TABLE 1.4 FDA-Approved Drugs with Their Companion Diagnostics Tests

Drug	Gene	CDx	Developer	Indication
Trastuzumab	Her2	Inform Her2/Neu	Roche (Ventana Medical Systems)	Breast carcinoma
		PathVysion Her2 DNA Probe Kit	Abbott (Abbott Molecular)	
		Pathway ANTI-Her2/NEU (4B5)	Roche (Ventana Medical Systems)	
		InSite Her2/NEU Kit	Biogenex Laboratories	
		SPOT-Light Her2 CISH Kit	Life Technologies (Thermo Fisher Scientific)	
		Bond Oracle Her2 IHC System	Leica Biosystems	
		Her2 CISH PharmDx Kit	Agilent Technologies (Dako)	
		INFORM Her2 DUAL ISH DNA Probe Cocktail	Roche (Ventana Medical Systems)	
		Herceptest™	Agilent Technologies (Dako)	
		Her2 FISH PharmDx Kit	Agilent Technologies (Dako)	
Cetuximab, Panitumumab	EGFR	DAKO EGFR PharmDx Kit	Agilent Technologies (Dako)	Colorectal cancer
Cetuximab	KRAS	Therascreen KRAS RGQ PCR Kit	Qiagen	
Imatinib mesylate	c-kit protein/CD 117	DAKO C-KIT PharmDx	Agilent Technologies (Dako)	Gastrointestinal stromal tumors (GIST)
Crizotinib	ALK	VYSIS ALK Break Apart FISH Probe Kit	Abbott (Abbott Molecular)	Non-small-cell lung cancer (NSCLC)
Erlotinib	EGFR	COBAS EGFR Mutation Test	Roche Molecular Systems	
Afatinib	EGFR	Therascreen EGFR RGQ PCR Kit	Qiagen	
Vemurafenib	BRAF	COBAS 4800 BRAF V600 Mutation Test	Roche Molecular Systems	Melanoma
Dabrafenib, Tramatenib	BRAF	THxID BRAF	bioMérieux	

The criteria for evaluating genetic tests are summarized by the four components of the ACCE analytic framework: analytic validity, clinical validity, clinical utility, and associated ethical, legal, and social implications [108]. Clinical validity refers to a test's ability to detect or predict the clinical disorder or phenotype associated with the genotype. Because most drug response phenotypes are multifactorial, it is not always easy to achieve the high clinical validity for pharmacogenetic testing that is often achieved with molecular testing for Mendelian disorders. Consequently, the positive predictive value of many pharmacogenetic assays can be low. Clinical utility of a test is a widely used measure of its usefulness in the clinic and of the changes in health outcomes that result. (See Figure 1.5.) However, given the multidimensional nature of this kind of measurement, there is rarely consensus as to its precise definition or how to adequately demonstrate it, particularly with regard to personalized medicine and pharmacogenetics [109]. Moreover, differences in regulators' requirements for clinical evidence for diagnostics vary between regions. The current FDA approach to diagnostics is generally more rigorous in terms of the clinical evidence required, while in Europe diagnostics manufacturers must satisfy essential safety and analytical performance requirements but are not required to establish that the diagnostic has an impact on clinical outcomes to obtain a CE marking [110].

The common benchmark for interventional evidence in medicine is an RCT (see Chapter 15), yet these are often resource-prohibitive for testing pharmacogenetic hypotheses and may be unethical regarding strong associations of severe adverse effects associated with high-risk genotypes. There is growing debate over and interest in the most appropriate study design that would provide evidence to accelerate translation of pharmacogenomic discoveries [109,111,112]. A key challenge in deciding which approach is most appropriate to establish the clinical utility of a new biomarker, is that genetic or other biomarkers often explain only a fraction of the observed outcome; other factors, including environment account for the remaining. RCTs, by definition, control all factors that could influence outcome, resulting in an artificially increased level of the standard of care. Thus they may dilute the effect of using a biomarker to guide decision making. Adaptive clinical trial designs are a modification of the standard RCT design for targeted therapeutics. The European Medicines Agency (EMA) in 2007and the FDA in 2010 published guidance on appropriate adaptive-trial conduct [113]. The use of retrospective and prospective case–control and cohort studies based on large databases has advantages over RCTs. The recent FDA updated drug label for carbamazepine regarding possible susceptibility to severe cutaneous reaction among Asian-ancestry carriers of the HLA-B*15:02 allele was based on retrospective observational data. Because of the low specificity of the test (i.e., most HLA-B*15:02 carriers will not develop an adverse event), it would be impossible (aside from being highly unethical) to study this marker in a RCT [112]. Abacavir-associated hypersensitivity reaction was observed in the drug development and post-approval stages, but it was a combination of retrospective and prospective studies that led to the identification of HLA-B*57:01 as a significant risk factor for this condition [114]. Subsequent RCTs showed that the positive predictive value of HLA-B*57:01 carriage for the development of abacavir-associated hypersensitivity reaction was ~50%, suggesting that the HLA-B*57:01 allotype is necessary but not sufficient for it [115]. Retrospective analyses have also proven useful in defining distinct patient subgroups that benefit preferentially from targeted therapeutic approaches [116–118].

1.7 ECONOMIC CHALLENGES

Healthcare expenditure in developed nations has been increasing faster than gross domestic product (GDP) over the past 40 years, mainly because of increasing per capita costs [119]. However, despite the success stories, pharmacogenomics has had limited impact on clinical practice to date. The list of FDA-approved companion diagnostics is still relatively small, with only 10% of currently marketed drugs having any pharmacogenomic information on the label, and the FDA currently mandates testing only for trastuzumab, cetuximab, and panitumumab.

There are significant clinical, financial, and ethical barriers to the successful implementation of pharmacogenomics. A significant trend in healthcare discussions is the focus on clinical utility and comparative effectiveness. The increasing emphasis on demonstrated therapeutic utility—"personal utility" in particular—will continue to have significant impact on clinical development strategies. At present, the extent of scientific knowledge related to the average safety and efficacy of potential drugs is typically much greater than knowledge about the underlying structure of drug responses in patient subpopulations and knowledge about the existence of predictive biomarkers to identify the most appropriate patient subpopulations for treatment.

However, utility can mean different things to the various stakeholders in the healthcare continuum. Traditionally, clinical development establishes the benefit–risk profile in order to support registration of a new treatment. Today, there is increased pressure to use clinical trials to produce evidence of relative efficacy and comparative effectiveness as compared with alternative courses of therapy [120]. Even with FDA-approved biomarker tests that appear to be clinically useful, lingering questions persist regarding their linkage to health outcomes [121].

Another noteworthy development is the role of healthcare payers in promoting evidence gathering to support the use of pharmacogenomics to inform treatment decisions [122]. Maintaining incentives for all stakeholders is not straightforward:

> *"Reimbursement frameworks that are dynamically linked, both initially and post-approval, to the patient population that benefits, rather than the patient population that is treated, provide incentives for co-developers to seek patient subpopulations that respond more effectively to treatment. Conversely, perpetuating the current reimbursement frameworks that are linked to existing, lower-efficacy treatment and diagnostic regimens discourages the development of stratified medicine" [123].*

According to a report issued by the Academy of Medical Sciences [110], for stratified medicine to become the standard in disease management the following requirements must be fulfilled:

Continued research to understand the genetic and molecular basis of disease.
Development and use of increasingly sophisticated and powerful informatics technology.
Improvement and standardization of clinical data collection and linkage with genomic and other databases.
Increased collection of tissues for biomarker research and evaluation, and their organization in national and international biobanks.
Greater efficiency and productivity in the development of therapeutics and diagnostics.
Flexible and novel approaches to regulatory assessments of innovative stratified medicine products.
Improved flexibility in pricing for stratified medicine products—for both diagnostic and associated therapy—to ensure cost-effectiveness for payers and encouragement of innovation.

Large, accessible datasets of patient information that employ the appropriate safeguards are essential to

BOX 1.3 Key Factors Affecting the Economic Value of Stratified Medicine

It is unclear whether a stratified medicine approach will enhance or reduce the economic value of a particular drug. After the introduction of a new therapeutic and its companion diagnostic, it is likely that a new equilibrium will initially be established that reduces patient populations because of diagnostic exclusion, enhanced efficacy, and safety. Market share will increase as the therapy becomes the preferred treatment [92]. In time, however, further market expansion will occur as underserved patients enter the market encouraged by the greater certainty of outcome should they qualify, and as long-term compliance improves [92] (see Figure 1.7).

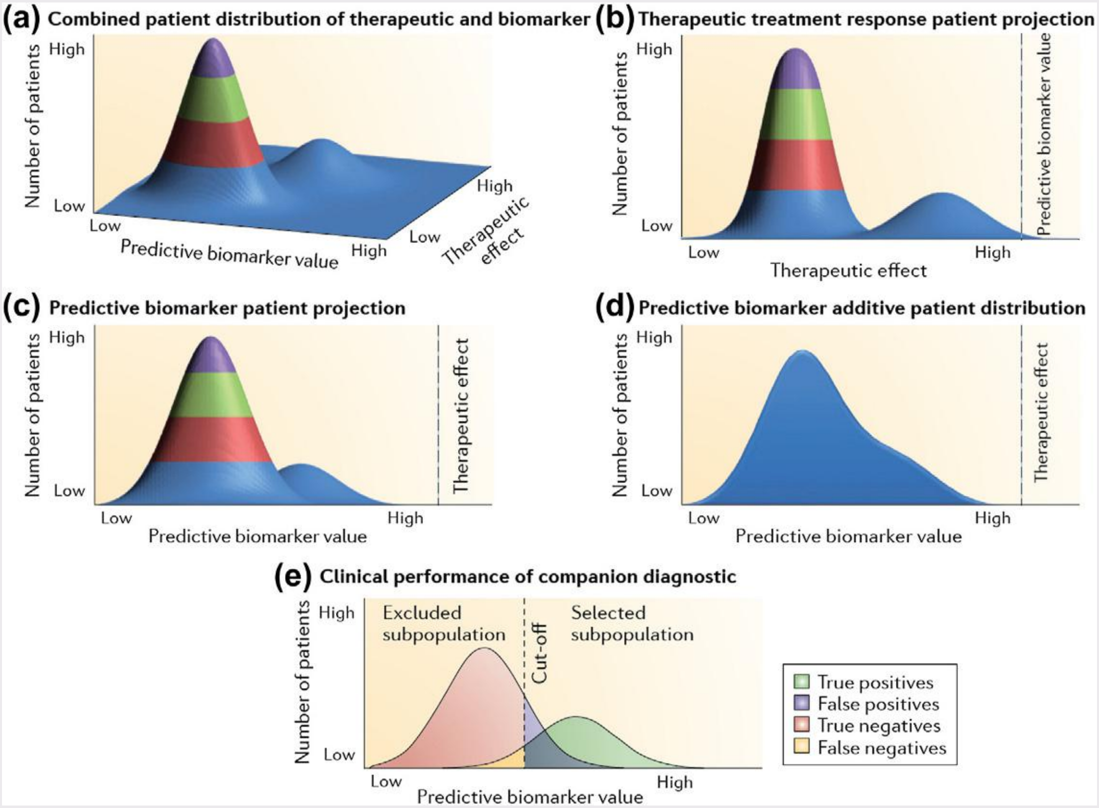


FIGURE 1.7 Key factors affecting the economic value of stratified medicine. Therapeutic effect is defined as the observed benefit of the treatment in patients. For traditional medicine this is usually simply stated as the mean benefit observed in the treated population compared to a placebo-treated or comparator-treated population. For stratified medicine this becomes more refined, as the treated population may refer to an all-comers–treated population, a biomarker-selected–treated population, or the underlying true treatment responders. The prevalence of the predictive biomarker is the fraction of the patient population that is selected by it and obviously has a direct and substantial effect on the number of patients who are subsequently treated. Notably, the selected population may be the converse of the patient population possessing the actual analyte measured. In addition, for a predictive biomarker with a continuous value, prevalence depends on the cut-off value selected.The clinical performance of the companion diagnostic is the ability of the test developed for the predictive biomarker (the companion diagnostic) to distinguish treatment responders from nonresponders. It is related to the combination of a test's sensitivity and specificity as well as negative and positive predictive values. It is important to recognize that most companion diagnostics perform substantially below perfect separation (with 100% values for all four metrics). The clinical performance of the companion diagnostic not only directly affects the number of patients who are potentially eligible for treatment but also affects the net benefit enrichment achieved, as patients who are selected by the companion diagnostic and are nonresponders also receive treatment, thereby reducing the observed average response. This has an impact on the price, market share, and speed of clinical adoption of the treatment.The patient population distributions for the therapeutic and the predictive biomarker usually separate imperfectly, as shown in part (a). Characterizing these distributions and their intersections provides much of the power and presents much of the development challenge of a stratified medicine approach. (a) Hypothetical patient distribution plotted against both therapeutic effect and a continuous-value predictive biomarker, illustrating a 30% high-responder subpopulation and a 70% low- or nonresponder subpopulation. (b) Projection of therapeutic effect only, showing a clean bimodal patient separation. (c) Two-dimensional projection of only the predictive biomarker value, with the therapeutic responder status preserved, allowing a nonadditive plot. Notably, two overlapping distributions result because clinical performance is less than perfect. (d) Observed patient distribution with a predictive biomarker, with no information on therapeutic effect. (e) Dichotomization of a continuous predictive biomarker into a companion diagnostic test to select a patient subpopulation for treatment. Both treatment responders and nonresponders scoring higher than the cut-off value (indicated by the vertical line) are selected for treatment. This example shows a case in which 30% of the total patient population responds to treatment but the chosen cut-off value generates a companion diagnostic with 90% sensitivity, 85% specificity, 96% negative predictive value, and 73% positive predictive value that selects 38% of the patients; of these, 27%will respond and 11% will not. *Source: Reprinted with permission from Macmillan Publishers Ltd: Nature Reviews Drug Discovery. Trusheim et al.* [123].

BOX 1.4 Potential Effects of Stratified Medicine on The Economic Value of Therapies (see Figure 1.8).

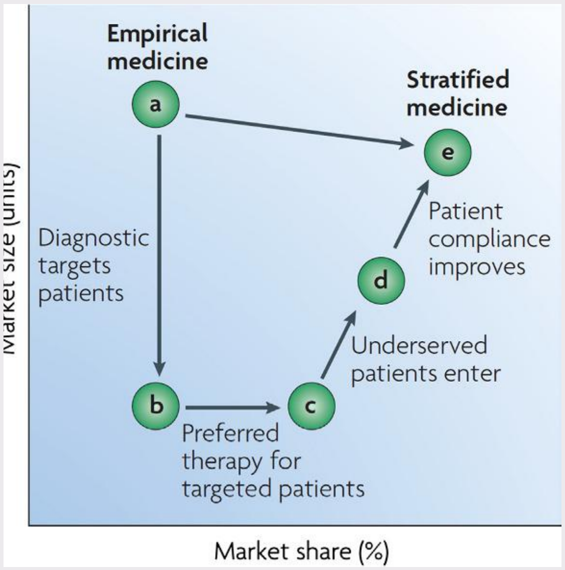

Empirical medicine
Stratified medicine
a
e
b
c
d
Diagnostic targets patients
Patient compliance improves
Underserved patients enter
Preferred therapy for targeted patients
Market size (units)
Market share (%)

FIGURE 1.8 Potential effects of stratified medicine on the economic value of therapies. Use of a diagnostic reduces the size of the treated patient population and thereby the potential market, but enhanced efficacy/safety increases market share as the therapy becomes the preferred treatment. Further market expansion occurs as underserved patients enter the market encouraged by the greater certainty of outcome should they qualify. Improved long-term compliance through diagnostic feedback and improved tolerability can further increase market size. *Source: Reprinted with permission from Macmillan Publishers Ltd: Nature Reviews Drug Discovery. Trusheim et al.* [92].

developing a stratified medicine program. They are central to the following areas: research to understand the molecular basis of disease; development of targeted interventions; effective regulation, health technology assessment, and evaluation of stratified medicine products; and stratification of treatment by physicians [110]. Co-developing a drug with a diagnostic to create a stratified medicine presents major barriers for product developers, regulators, payers, and physicians.

In reality, because of the different development timescales involved, there is little incentive for diagnostic manufacturers to co-develop a diagnostic when a drug enters its preclinical or clinical phase. On average, it takes 9–12 years to develop a drug in contrast to 3–5 years for a diagnostic. This problem is compounded by the high risk that a drug may fail to meet safety and efficacy requirements, making the associated diagnostic redundant. From an industry perspective, there is a financial "double disincentive" to stratify medicines, as manufacturers stand to pay more for R&D and see their market size shrink.

Trusheim et al. [123] performed an analysis using data from case studies in oncology and Alzheimer's disease, coupled with integrated computational modeling of clinical outcomes and developer economic value, to quantify the effects of decisions related to key issues such as the design of clinical trials. Out of the over 90 factors input into the economic models, three in particular drive economic value for stratified medicine compared to empirical medicine: relative therapeutic performance, the prevalence of the biomarker, and the biomarker's clinical performance. Enhanced therapeutic effects not only benefit patients but also enable product differentiation, market share gains, and pricing justification; each of these benefits—alone and especially in combination—return economic value to the developers of the drug and the diagnostic.

1.8 CONCLUSIONS

Clearly there is great potential for pharmacogenomics to improve the benefit–risk profile for individual patients and to improve the effectiveness of drug development. The growing number of drug labels with pharmacogenetic information, more treatment guideline recommendations based on genetics, and escalating investment in pharmacogenetics by industry, academia, governments, and institutions are all evidence that genomics is delivering on its promise, albeit slowly. There is burgeoning interest in implementing pharmacogenomic discoveries in clinical practice, and this requires cooperation from all stakeholders to develop the evidence path necessary to define utility.

Next-generation sequencing technology will make economically viable the sequencing of the entire genome of an individual (or his or her tumor) to create a personal variant file. The whole genome could ultimately function as a set of biomarkers, with sequencing of individual tumor genomes used to inform treatment decisions based on in silico models [110,124]. In the future, prescribing may depend on matching the drug to the patient using his or her stored genetic profile coupled with an improved understanding of the molecular pathway of disease, thus reducing the need for traditional companion diagnostics.

REFERENCES

[1] Spear BB, Heath-Chiozzi M, Huff J. Clinical application of pharmacogenetics. Trends Mol Med 2001;7(5):201–4.

[2] Regierer B, Zazzu V, Sudbrak R, Kuhn A, Lehrach H. Future of medicine: models in predictive diagnostics and personalized medicine. Adv Biochem Eng Biotechnol 2013;133:15–33.

[3] Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 2004;329 (7456):15–19.

[4] Budnitz DS, Shehab N, Kegler SR, Richards CL. Medication use leading to emergency department visits for adverse drug events in older adults. Ann Intern Med 2007;147(11):755–65.

[5] Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;279(15):1200–5.

[6] Miguel A, Azevedo LF, Araujo M, Pereira AC. Frequency of adverse drug reactions in hospitalized patients: a systematic review and meta-analysis. Pharmacoepidemiol Drug Saf 2012;21(11):1139–54.

[7] Burns DK. Developing pharmacogenetic evidence throughout clinical development. Clin Pharmacol Ther 2010;88(6):867–70.

[8] Vesell ES. Pharmacogenetic perspectives gained from twin and family studies. Pharmacol Ther 1989;41(3):535–52.

[9] Kalow W, Tang BK, Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. Pharmacogenetics 1998;8(4):283–9.

[10] Alving AS, Carson PE, Flanagan CL, Ickes CE. Enzymatic deficiency in primaquine-sensitive erythrocytes. Science 1956;124(3220):484–5.

[11] Hughes HB, Biehl JP, Jones AP, Schmidt LH. Metabolism of isoniazid in man as related to the occurrence of peripheral neuritis. Am Rev Tuberc 1954;70(2):266–73.

[12] Lehmann H, Ryan E. The familial incidence of low pseudocholinesterase level. Lancet 1956;271(6934):124.

[13] Gonzalez FJ, Skoda RC, Kimura S, Umeno M, Zanger UM, Nebert DW, et al. Characterization of the common genetic defect in humans deficient in debrisoquine metabolism. Nature 1988;331(6155):442–6.

[14] Evans WE, McLeod HL. Pharmacogenomics–drug disposition, drug targets, and side effects. N Engl J Med 2003;348(6):538–49.

[15] Weinshilboum R, Wang L. Pharmacogenomics: bench to bedside. Nat Rev Drug Discov 2004;3(9):739–48.

[16] FDA. Pharmacogenomic Biomarkers in Drug Labels. <http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm>. 2013.

[17] Jensen EV. Estrogen receptors in hormone-dependent breast cancers. Cancer Res 1975;35(11 Pt. 2):3362–4.

[18] Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365(9472):1687–717.

[19] Borges S, Desta Z, Li L, Skaar TC, Ward BA, Nguyen A, et al. Quantitative effect of CYP2D6 genotype and inhibitors on tamoxifen metabolism: implication for optimization of breast cancer treatment. Clin Pharmacol Ther 2006;80(1):61–74.

[20] Garte S, Gaspari L, Alexandrie AK, Ambrosone C, Autrup H, Autrup JL, et al. Metabolic gene polymorphism frequencies in control populations. Cancer Epidemiol Biomarkers Prev 2001;10(12):1239–48.

[21] Goetz MP, Knox SK, Suman VJ, Rae JM, Safgren SL, Ames MM, et al. The impact of cytochrome P450 2D6 metabolism in women receiving adjuvant tamoxifen. Breast Cancer Res Treat 2007;101(1):113–21.

[22] Schroth W, Goetz MP, Hamann U, Fasching PA, Schmidt M, Winter S, et al. Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. JAMA 2009;302(13):1429–36.

[23] Punglia RS, Burstein HJ, Winer EP, Weeks JC. Pharmacogenomic variation of CYP2D6 and the choice of optimal adjuvant endocrine therapy for postmenopausal breast cancer: a modeling analysis. J Natl Cancer Inst 2008;100(9):642–8.

[24] Andersson T, Holmberg J, Rohss K, Walan A. Pharmacokinetics and effect on caffeine metabolism of the proton pump inhibitors, omeprazole, lansoprazole, and pantoprazole. Br J Clin Pharmacol 1998;45(4):369–75.

[25] Pearce RE, Rodrigues AD, Goldstein JA, Parkinson A. Identification of the human P450 enzymes involved in lansoprazole metabolism. J Pharmacol Exp Ther 1996;277(2):805–16.

[26] Bottiger Y, Tybring G, Gotharson E, Bertilsson L. Inhibition of the sulfoxidation of omeprazole by ketoconazole in poor and extensive metabolizers of S-mephenytoin. Clin Pharmacol Ther 1997;62(4):384–91.

[27] de Morais SM, Wilkinson GR, Blaisdell J, Nakamura K, Meyer UA, Goldstein JA. The major genetic defect responsible for the polymorphism of S-mephenytoin metabolism in humans. J Biol Chem 1994;269(22):15419–22.

[28] Sim SC, Risinger C, Dahl ML, Aklillu E, Christensen M, Bertilsson L, et al. A common novel CYP2C19 gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. Clin Pharmacol Ther 2006;79(1):103–13.

[29] Furuta T, Shirai N, Xiao F, Ohashi K, Ishizaki T. Effect of high-dose lansoprazole on intragastic pH in subjects who are homozygous extensive metabolizers of cytochrome P4502C19. Clin Pharmacol Ther 2001;70(5):484–92.

[30] Cooper GM, Johnson JA, Langaee TY, Feng H, Stanaway IB, Schwarz UI, et al. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. Blood 2008;112(4):1022–7.

[31] Limdi NA, Beasley TM, Crowley MR, Goldstein JA, Rieder MJ, Flockhart DA, et al. VKORC1 polymorphisms, haplotypes and haplotype groups on warfarin dose among African-Americans and European-Americans. Pharmacogenomics 2008;9(10):1445–58.

[32] Takeuchi F, McGinnis R, Bourgeois S, Barnes C, Eriksson N, Soranzo N, et al. A genome-wide association study confirms

VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. PLoS Genet 2009;5(3):e1000433.

[33] Horne BD, Lenzini PA, Wadelius M, Jorgensen AL, Kimmel SE, Ridker PM, et al. Pharmacogenetic warfarin dose refinements remain significantly influenced by genetic factors after one week of therapy. Thromb Haemost 2012;107(2):232–40.

[34] Anderson JL, Horne BD, Stevens SM, Grove AS, Barton S, Nicholas ZP, et al. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. Circulation 2007;116(22):2563–70.

[35] Garcia DA. Warfarin and pharmacogenomic testing: the case for restraint. Clin Pharmacol Ther 2008;84(3):303–5.

[36] Yen T, Nightingale BN, Burns JC, Sullivan DR, Stewart PM. Butyrylcholinesterase (BCHE) genotyping for post-succinylcholine apnea in an Australian population. Clin Chem 2003;49(8):1297–308.

[37] Evans DA, Manley KA, McKusick VA. Genetic control of isoniazid metabolism in man. Br Med J 1960;2(5197):485–91.

[38] Perry Jr HM, Tan EM, Carmody S, Sakamoto A. Relationship of acetyl transferase activity to antinuclear antibodies and toxic symptoms in hypertensive patients treated with hydralazine. J Lab Clin Med 1970;76(1):114–25.

[39] Gravante G, Delogu D, Marianetti M, Trombetta M, Esposito G, Montone A. Toxic epidermal necrolysis and Steven Johnson syndrome: 11-years experience and outcome. Eur Rev Med Pharmacol Sci 2007;11(2):119–27.

[40] Mockenhaupt M. The current understanding of Stevens-Johnson syndrome and toxic epidermal necrolysis. Expert Rev Clin Immunol 2011;7(6):803–13.

[41] Chung WH, Hung SI, Hong HS, Hsih MS, Yang LC, Ho HC, et al. Medical genetics: a marker for Stevens-Johnson syndrome. Nature 2004;428(6982):486.

[42] Ferrell Jr PB, McLeod HL. Carbamazepine, HLA-B*1502 and risk of Stevens-Johnson syndrome and toxic epidermal necrolysis: US FDA recommendations. Pharmacogenomics 2008;9(10):1543–6.

[43] Tidd DM, Paterson AR. A biochemical mechanism for the delayed cytotoxic reaction of 6-mercaptopurine. Cancer Res 1974;34(4):738–46.

[44] Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. Am J Hum Genet 1980;32(5):651–62.

[45] Schaeffeler E, Fischer C, Brockmeier D, Wernet D, Moerike K, Eichelbaum M, et al. Comprehensive analysis of thiopurine S-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel TPMT variants. Pharmacogenetics 2004;14(7):407–17.

[46] Collie-Duguid ES, Pritchard SC, Powrie RH, Sludden J, Collier DA, Li T, et al. The frequency and distribution of thiopurine methyltransferase alleles in Caucasian and Asian populations. Pharmacogenetics 1999;9(1):37–42.

[47] Hon YY, Fessing MY, Pui CH, Relling MV, Krynetski EY, Evans WE. Polymorphism of the thiopurine S-methyltransferase gene in African-Americans. Hum Mol Genet 1999;8(2):371–6.

[48] Chang JG, Lee LS, Chen CM, Shih MC, Wu MC, Tsai FJ, et al. Molecular analysis of thiopurine S-methyltransferase alleles in South-east Asian populations. Pharmacogenetics 2002;12(3):191–5.

[49] Kumagai K, Hiyama K, Ishioka S, Sato H, Yamanishi Y, McLeod HL, et al. Allelotype frequency of the thiopurine methyltransferase (TPMT) gene in Japanese. Pharmacogenetics 2001;11(3):275–8.

[50] Iyer L, Hall D, Das S, Mortell MA, Ramirez J, Kim S, et al. Phenotype-genotype correlation of in vitro SN-38 (active metabolite of irinotecan) and bilirubin glucuronidation in human liver tissue with UGT1A1 promoter polymorphism. Clin Pharmacol Ther 1999;65(5):576–82.

[51] Premawardhena A, Fisher CA, Liu YT, Verma IC, de SS, Arambepola M, et al. The global distribution of length polymorphisms of the promoters of the glucuronosyltransferase 1 gene (UGT1A1): hematologic and evolutionary implications. Blood Cells Mol Dis 2003;31(1):98–101.

[52] Giacomini KM, Yee SW, Ratain MJ, Weinshilboum RM, Kamatani N, Nakamura Y. Pharmacogenomics and patient care: one size does not fit all. Sci Transl Med 2012;4(153) (153ps18).

[53] Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenomics J 2008;8(3):186–95.

[54] Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe'er I, Floratos A, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. Nat Genet 2009;41(7):816–9.

[55] Singer JB, Lewitzky S, Leroy E, Yang F, Zhao X, Klickstein L, et al. A genome-wide study identifies HLA alleles associated with lumiracoxib-related liver injury. Nat Genet 2010;42(8):711–4.

[56] Lucena MI, Molokhia M, Shen Y, Urban TJ, Aithal GP, Andrade RJ, et al. Susceptibility to amoxicillin-clavulanate-induced liver injury is influenced by multiple HLA class I and II alleles. Gastroenterology 2011;141(1):338–47.

[57] Ozeki T, Mushiroda T, Yowang A, Takahashi A, Kubo M, Shirakata Y, et al. Genome-wide association study identifies HLA-A*3101 allele as a genetic risk factor for carbamazepine-induced cutaneous adverse drug reactions in Japanese population. Hum Mol Genet 2011;20(5):1034–41.

[58] Chen P, Lin JJ, Lu CS, Ong CT, Hsieh PF, Yang CC, et al. Carbamazepine-induced toxic effects and HLA-B*1502 screening in Taiwan. N Engl J Med 2011;364(12):1126–33.

[59] Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy–a genomewide study. N Engl J Med 2008;359(8):789–99.

[60] Volpi S, Heaton C, Mack K, Hamilton JB, Lannan R, Wolfgang CD, et al. Whole genome association study identifies polymorphisms associated with QT prolongation during iloperidone treatment of schizophrenia. Mol Psychiatry 2009;14(11):1024–31.

[61] Sarasquete ME, Garcia-Sanz R, Marin L, Alcoceba M, Chillon MC, Balanzategui A, et al. Bisphosphonate-related osteonecrosis of the jaw is associated with polymorphisms of the cytochrome P450 CYP2C8 in multiple myeloma: a genome-wide single nucleotide polymorphism analysis. Blood 2008;112(7):2709–12.

[62] Fellay J, Thompson AJ, Ge D, Gumbs CE, Urban TJ, Shianna KV, et al. ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. Nature 2010;464(7287):405–8.

[63] Pare G, Eriksson N, Lehr T, Connolly S, Eikelboom J, Ezekowitz MD, et al. Genetic determinants of dabigatran plasma levels and their relation to bleeding. Circulation 2013;127(13):1404–12.

[64] Aberg K, Adkins DE, Liu Y, McClay JL, Bukszar J, Jia P, et al. Genome-wide association study of antipsychotic-induced QTc interval prolongation. Pharmacogenomics J 2012;12(2):165–72.

[65] Tanaka Y, Kurosaki M, Nishida N, Sugiyama M, Matsuura K, Sakamoto N, et al. Genome-wide association study identified ITPA/

DDRGK1 variants reflecting thrombocytopenia in pegylated interferon and ribavirin therapy for chronic hepatitis C. Hum Mol Genet 2011;20(17):3507–16.
[66] Tan XL, Moyer AM, Fridley BL, Schaid DJ, Niu N, Batzler AJ, et al. Genetic variation predicting cisplatin cytotoxicity associated with overall survival in lung cancer patients receiving platinum-based chemotherapy. Clin Cancer Res 2011;17(17):5801–11.
[67] Chantarangsu S, Mushiroda T, Mahasirimongkol S, Kiertiburanakul S, Sungkanuparph S, Manosuthi W, et al. Genome-wide association study identifies variations in 6p21.3 associated with nevirapine-induced rash. Clin Infect Dis 2011;53(4):341–8.
[68] Ingle JN, Schaid DJ, Goss PE, Liu M, Mushiroda T, Chapman JA, et al. Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. J Clin Oncol 2010;28(31):4674–82.
[69] Huang RS, Johnatty SE, Gamazon ER, Im HK, Ziliak D, Duan S, et al. Platinum sensitivity-related germline polymorphism discovered via a cell-based approach and analysis of its association with outcome in ovarian cancer patients. Clin Cancer Res 2011;17(16):5490–500.
[70] Drago A, Crisafulli C, Serretti A. The genetics of antipsychotic induced tremors: a genome-wide pathway analysis on the STEP-BD SCP sample. Am J Med Genet B Neuropsychiatr Genet 2011;156B(8):975–86.
[71] Adkins DE, Clark SL, Aberg K, Hettema JM, Bukszar J, McClay JL, et al. Genome-wide pharmacogenomic study of citalopram-induced side effects in STAR*D. Transl Psychiatry 2012;2:e129.
[72] Srinivasan Y, Sasa M, Honda J, Takahashi A, Uno S, Kamatani N, et al. Genome-wide association study of epirubicin-induced leukopenia in Japanese patients. Pharmacogenet Genomics 2011;21(9):552–8.
[73] Tohkin M, Kaniwa N, Saito Y, Sugiyama E, Kurose K, Nishikawa J, et al. A whole-genome association study of major determinants for allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis in Japanese patients. Pharmacogenomics J 2013;13(1):60–9.
[74] Kim JH, Park BL, Cheong HS, Bae JS, Park JS, Jang AS, et al. Genome-wide and follow-up studies identify CEP68 gene variants associated with risk of aspirin-intolerant asthma. PLoS One 2010;5(11):e13818.
[75] Innocenti F, Owzar K, Cox NL, Evans P, Kubo M, Zembutsu H, et al. A genome-wide association study of overall survival in pancreatic cancer patients treated with gemcitabine in CALGB 80303. Clin Cancer Res 2012;18(2):577–84.
[76] Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, et al. A genomewide association study of citalopram response in major depressive disorder. Biol Psychiatry 2010;67(2):133–8.
[77] Chung CM, Wang RY, Chen JW, Fann CS, Leu HB, Ho HY, et al. A genome-wide association study identifies new loci for ACE activity: potential implications for response to ACE inhibitor. Pharmacogenomics J 2010;10(6):537–44.
[78] Turner ST, Bailey KR, Schwartz GL, Chapman AB, Chai HS, Boerwinkle E. Genomic association analysis identifies multiple loci influencing antihypertensive response to an angiotensin II receptor blocker. Hypertension 2012;59(6):1204–11.
[79] Turner ST, Boerwinkle E, O'Connell JR, Bailey KR, Gong Y, Chapman AB, et al. Genomic association analysis of common variants influencing antihypertensive response to hydrochlorothiazide. Hypertension 2013;62(2):391–7.
[80] Tantisira KG, Lasky-Su J, Harada M, Murphy A, Litonjua AA, Himes BE, et al. Genomewide association between GLCCI1 and response to glucocorticoid therapy in asthma. N Engl J Med 2011;365(13):1173–83.
[81] Tantisira KG, Damask A, Szefler SJ, Schuemann B, Markezich A, Su J, et al. Genome-wide association identifies the T gene as a novel asthma pharmacogenetic locus. Am J Respir Crit Care Med 2012;185(12):1286–91.
[82] Shuldiner AR, O'Connell JR, Bliden KP, Gandhi A, Ryan K, Horenstein RB, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. JAMA 2009;302(8):849–57.
[83] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461(7262):399–401.
[84] Zhou K, Bellenguez C, Spencer CC, Bennett AJ, Coleman RL, Tavendale R, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. Nat Genet 2011;43(2):117–20.
[85] Trevino LR, Shimasaki N, Yang W, Panetta JC, Cheng C, Pei D, et al. Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. J Clin Oncol 2009;27(35):5972–8.
[86] Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 2004;3(8):711–5.
[87] Randazzo M, Terness P, Opelz G, Kleist C. Active-specific immunotherapy of human cancers with the heat shock protein Gp96-revisited. Int J Cancer 2012;130(10):2219–31.
[88] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126(4):663–76.
[89] Ferreira LM, Mostajo-Radji MA. How induced pluripotent stem cells are redefining personalized medicine. Gene 2013;520(1):1–6.
[90] Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA. Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. Stat Med 2012;31(17):1887–902.
[91] Von KM, Moore JC. Stepped care for back pain: activating approaches for primary care. Ann Intern Med 2001;134(9 Pt 2):911–7.
[92] Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. Nat Rev Drug Discov 2007;6(4):287–93.
[93] Hu SX, Foster T, Kieffaber A. Pharmacogenomics and personalized medicine: mapping of future value creation. Biotechniques 2005;39(10 Suppl):S1–S6.
[94] Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet 2011;378(9802):1560–71.
[95] Burstein HJ. The distinctive nature of HER2-positive breast cancers. N Engl J Med 2005;353(16):1652–4.
[96] Brown MJ. Blood-pressure-lowering treatment. Lancet 2001;357(9257):715–6.
[97] Padmanabhan S, Wallace C, Munroe PB, Dobson R, Brown M, Samani N, et al. Chromosome 2p shows significant linkage to antihypertensive response in the British Genetics of Hypertension Study. Hypertension 2006;47(3):603–8.
[98] Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med 2010;363(4):301–4. July 22

[99] Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clin Pharmacol Ther 2001; 69(3):89–95.

[100] Buyse M, Sargent DJ, Grothey A, Matheson A, de GA. Biomarkers and surrogate end points–the challenge of statistical validation. Nat Rev Clin Oncol 2010;7(6):309–17.

[101] Frank R, Hargreaves R. Clinical biomarkers in drug discovery and development. Nat Rev Drug Discov 2003;2(7):566–80.

[102] Buyse M, Michiels S, Sargent DJ, Grothey A, Matheson A, de GA. Integrating biomarkers in clinical trials. Expert Rev Mol Diagn 2011;11(2):171–82.

[103] Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. EPMA J 2013;4(1):7.

[104] Aspinall MG, Hamermesh RG. Realizing the promise of personalized medicine. Harv Bus Rev 2007;85(10):108–17. 165

[105] Phillips KA, Van BS, Issa AM. Diagnostics and biomarker development: priming the pipeline. Nat Rev Drug Discov 2006;5(6):463–9.

[106] Schubert C. Cancer drugs find a companion with new diagnostic tests. Nat Med 2011;17(10):1157.

[107] Love D, Stratton E, Stocum M. Best practices for companion diagnostic and therapeutic development: translating between the stakeholders. N Biotechnol 2012;29(6):689–94.

[108] Sanderson S, Zimmern R, Kroese M, Higgins J, Patch C, Emery J. How can the evaluation of genetic tests be enhanced? Lessons learned from the ACCE framework and evaluating genetic tests in the United Kingdom. Genet Med 2005;7(7):495–500.

[109] Lesko LJ, Zineh I, Huang SM. What is clinical utility and why should we care? Clin Pharmacol Ther 2010;88(6):729–33.

[110] Academy of Medical Sciences, Realising the potential of stratified medicine. <http://www.acmedsci.ac.uk/download.php?file=/images/pressRelease/FINALFOR.pdf>. 2013.

[111] Teagarden JR. Warfarin and pharmacogenomic testing: what would Pascal do? Pharmacotherapy 2009;29(3):245–7.

[112] Frueh FW. Back to the future: why randomized controlled trials cannot be the answer to pharmacogenomics and personalized medicine. Pharmacogenomics 2009;10(7):1077–81.

[113] Editorial. Time to adapt. Nature 2010;464(7293):1245–6.

[114] Hughes AR, Spreen WR, Mosteller M, Warren LL, Lai EH, Brothers CH, et al. Pharmacogenetics of hypersensitivity to abacavir: from PGx hypothesis to confirmation to clinical utility. Pharmacogenomics J 2008;8(6):365–74.

[115] Mallal S, Phillips E, Carosi G, Molina JM, Workman C, Tomazic J, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008;358(6):568–79.

[116] Amado RG, Wolf M, Peeters M, Van CE, Siena S, Freeman DJ, et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. J Clin Oncol 2008;26(10):1626–34.

[117] Hayes DF, Thor AD, Dressler LG, Weaver D, Edgerton S, Cowan D, et al. HER2 and response to paclitaxel in node-positive breast cancer. N Engl J Med 2007;357(15):1496–506.

[118] Takano T, Fukui T, Ohe Y, Tsuta K, Yamamoto S, Nokihara H, et al. EGFR mutations predict survival benefit from gefitinib in patients with advanced lung adenocarcinoma: a historical comparison of patients treated before and after gefitinib approval in Japan. J Clin Oncol 2008;26(34):5589–95.

[119] US National Bureau of Economic Researc. Who's going broke? Comparing growth in healthcare costs in ten OECD countries. <http://www.nber.org/papers/w11833>. 2005.

[120] Eichler HG, Bloechl-Daum B, Abadie E, Barnett D, Konig F, Pearson S. Relative efficacy of drugs: an emerging issue between regulatory agencies and third-party payers. Nat Rev Drug Discov 2010;9(4):277–91.

[121] Lesko LJ, Woodcock J. Translation of pharmacogenomics and pharmacogenetics: a regulatory perspective. Nat Rev Drug Discov 2004;3(9):763–9.

[122] Epstein RS, Frueh FW, Geren D, Hummer D, McKibbin S, O'Connor S, et al. Payer perspectives on pharmacogenomics testing and drug development. Pharmacogenomics 2009;10(1):149–51.

[123] Trusheim MR, Burgess B, Hu SX, Long T, Averbuch SD, Flynn AA, et al. Quantifying factors for the success of stratified medicine. Nat Rev Drug Discov 2011;10(11):817–33.

[124] European Science Foundation. Personalised medicine for the European citizen: towards more precise medicine for the diagnosis, treatment and prevention of disease (iPM). <http://www.esf.org/uploads/media/Personalised_Medicine.pdf>. 2012.

This page is intentionally left blank

Part II

Basic Genetics

This page is intentionally left blank

Chapter 2

Basic Genetics: The Cell, Mitosis and Meiosis, and Mendelian Laws

Guan Wang
Institute of Cardiovascular and Medical Sciences, College of Medicine, Veterinary and Life Sciences, University of Glasgow, UK

Chapter Outline

2.1 Overview of the Cell: Anatomy, Components, and Function 29
  2.1.1 Nucleus and Chromosomes 29
  2.1.2 Membrane-Bound Organelles in Eukaryotic Cells 30
  2.1.3 Nonmembranous Organelles in Eukaryotic Cells 30
2.2 Cell Reproduction: Cell Cycle and Mitosis 30
2.3 Meiosis 33
  2.3.1 Genetic Variation Introduced by Meiosis: The Mechanisms 34
2.4 Mendelian Laws 35
  2.4.1 The Law of Segregation 35
  2.4.2 The Law of Independent Assortment 37
  2.4.3 Mendelian Patterns of Inheritance in Humans 37
2.5 Public Databases for Biomedical Research in Humans 38
2.6 Conclusion 39
Glossary 39
Acknowledgments 40
References 40

2.1 OVERVIEW OF THE CELL: ANATOMY, COMPONENTS, AND FUNCTION

Each cell is a basic building block for any organism living on Earth. These organisms are either unicellular or multicellular. In a unicellular organism, the cell or the organism performs all necessary activities required to sustain life. In a multicellular organism, various cells with specialized functions cooperate and facilitate proper functioning of the whole organism; nevertheless, each single cell in a multicellular organism is seen as the organism's basic unit of structure and function. In other words, life fundamentally occurs at the cellular level.

There are two kinds of living organisms: prokaryotes and eukaryotes. The main difference between the two is that eukaryotes have a well-defined nucleus enclosed in a membranous nuclear envelope; prokaryotes have no nucleus but a region called the nucleoid not separated from the rest of the cell. In eukaryotic cells, the region between the nucleus and the cell membrane, called the cytoplasm, is where the membrane-enclosed organelles are suspended in a semifluid medium called the cytosol. Unlike eukaryotic cells, prokaryotic cells lack most of the organelles in the cytoplasm. Prokaryotic cells are found only in the kingdom Monera (such as bacteria), while protists, fungi, plants, and animals are all eukaryotes. A detailed comparison of eukaryotic and prokaryotic cells is given in Table 2.1. The discussion of cell structure and organelle function in this chapter applies to eukaryotes.

2.1.1 Nucleus and Chromosomes

The nuclear envelope separates the two major cellular compartments: the nucleoplasm and the cytoplasm. The DNA and its associated proteins are referred to as chromatin and reside in the nucleus in eukaryotes. The chromosomes carrying the cell's genes and the machinery to express them reside inside the nucleus. Interactions with histones and other proteins fold each chromosome compactly enough to fit inside the nucleus. In higher organisms, each somatic cell contains

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00002-5
© 2014 Elsevier Inc. All rights reserved.

TABLE 2.1 Comparison of Eukaryotic and Prokaryotic Cells

Feature	Eukaryote	Prokaryote
Cell size	Typically 10–100 μm	Typically 1–10 μm
Nucleus	Present and membrane-bound	Absent but has a nucleoid region without a nuclear membrane
Other organelles	Many membrane-bound organelles	No membrane-bound ribosomes
Genetic material (DNA)	Located in the nucleus and surrounded by a nuclear membrane	Concentrated in the nucleoid
DNA arrangement	Multiple linear chromosomes	Single circular chromosome
RNA synthesis	In the nucleus	In the cytoplasm
Protein synthesis	In the cytoplasm	In the cytoplasm
Cytoplasmic structure	Cytoskeleton and cytoplasmic streaming	No cytoskeleton or cytoplasmic streaming
Organization	Often multicellular	Usually unicellular
Cell division	Mitosis and meiosis	Fission or budding

one set of chromosomes inherited from the maternal parent and a comparable set of chromosomes (homologous, or homologues) from the paternal parent. The number of chromosomes in this dual set is the diploid number. Sex cells, or gametes, which contain half the number of chromosome sets found in somatic cells, are referred to as haploid cells.

2.1.2 Membrane-Bound Organelles in Eukaryotic Cells

The endomembrane system separates the cell into different compartments, or organelles, such as the nucleus, the endoplasmic reticulum (ER), the Golgi apparatus, and lysosomes (see Table 2.2). The endomembrane system is derived from the ER and flows to the Golgi apparatus, from which lysosomes bud. The ER is a continuous system of flattened membrane sacks and tubules that is specialized for protein processing and lipid biosynthesis. The endomembrane system is important for the cell's compartmental organization to function independently and properly. Ribosomes that synthesize proteins destined for insertion into cellular membranes or for export from the cell associate with specialized regions of the ER, called the rough ER owing to the attached ribosomes.

For the most common organelles in eukaryotic cells, the structure and function of each are illustrated, using an animal cell as an example, in Figure 2.1 and Table 2.2.

2.1.3 Nonmembranous Organelles in Eukaryotic Cells

The shape of a cell is maintained by a network of fibers throughout the cytoplasm, called the cytoskeleton. In addition to structural support, the cytoskeleton functions in cell motility, such as the movement of organelles within the cell. The three fiber components of the cytoskeleton are microtubules, microfilaments, and intermediate filaments.

Microtubules are straight, hollow tubes made of tubulins. Two tubulin types, α-tubulin and β-tubulin, are similar molecular units bonded together to form the microtubule wall. Microtubules support the cell and guide the movement of organelles by providing tracks for them equipped with specialized proteins known as motor molecules. Microtubules also help in the separation of chromosomes during cell division, as discussed in later sections.

Microfilaments are solid rods built from actin molecules. Actin molecules are present in chains twisted in a helix form. Microfilaments are especially concentrated and well ordered in muscle cells and are involved in muscle cell contraction. They also play a role in supporting the cell along with the rest of the cytoskeleton.

Intermediate filaments are smaller than microtubules but larger than microfilaments in diameter. These filaments can remain in their original shape, while the microtubules and microfilaments are often disassembled and reconstructed in different parts of the cytoplasm. This suggests that they are especially important in maintaining the shape of a cell and fixing the position of certain organelles.

2.2 CELL REPRODUCTION: CELL CYCLE AND MITOSIS

Cell division may result in cell reproduction in a unicellular organism, cell growth, and development of multicellular organisms, as well as the renewal and repair of tissues in mature organisms. In a prokaryotic cell (e.g., bacterium), the type of cell division is called binary fission. Bacterial DNA is attached to the cell membrane. In preparation for

TABLE 2.2 Most Common Characteristics of Membrane-Bound Organelles in Animal Eukaryotic Cells

Organelle	Feature	Function
Nucleus	• Nuclear envelope: lipid double layer • Nuclear pores • Chromatin (mass of DNA)/chromosome (DNA double helix, visible during cell division) • Nucleolus	• Separates the nucleus's contents from the cytoplasm • Regulate nuclear transportation • Holds the cell's genetic material; human somatic cell: 46 chromosomes; sex cells: 23 chromosomes • Synthesizes the ribosome's components to be assembled in the cytoplasm
Ribosomes	• Free ribosomes suspended in cytosol • Ribosomes attached to the endoplasmic reticulum • Two ribosome types are structurally identical and functionally interchangeable	• Synthesize proteins to function within the cytosol • Synthesize proteins to be included in membranes; pack certain organelles (e.g., lysosomes) or export from cell • Interchange occurs when the cell's metabolism changes
Endoplasmic reticulum (ER)	• Rough ER (with studded ribosomes) is confluent with the nuclear outer membrane • Smooth ER (lacking ribosomes) is connected to the rough ER	• Wraps secretory proteins in the transport vesicles (e.g., glycoproteins) and produces membranes for the ER itself or other components of the endomembrane system by adding membrane proteins and phospholipids • Cell type-dependent and may function in lipid synthesis, carbohydrate metabolism, and drug detoxification.
Golgi apparatus	• Stacks of flattened membranous sacs • *cis* face[a] • *trans* face[b]	• Receives, modifies, stores, and dispatches transport vesicles (including their contents, e.g., secretory proteins from the ER) • Receives vesicles • Dispatches vesicles (ER product modification occurs during transit from the *cis* face to the *trans* face)
Lysosomes	• Membrane-bound sacs containing hydrolytic enzymes	• Digest macromolecules through phagocytosis or autophagy
Peroxisomes	• Single membrane-bound organelles containing a collection of enzymes	• Transfer hydrogen from various substrates to oxygen, producing hydrogen peroxide as a by-product. These reactions serve various metabolic functions—e.g., breakdown of fatty acids for mitochondria respiration and detoxification of alcohol in the liver
Mitochondria	• Double-membrane-bound organelles containing ribosomes and DNA • Infoldings of the inner membrane form the cristae • Number of mitochondria varies with the metabolic activity of a cell	• Convert energy for cellular respiration

[a] *One of the two poles of the Golgi apparatus located near ER and receiving the content carried by transport vesicles from ER.* [b] *The other pole of the Golgi apparatus, giving rise to vesicles that travel to other sites.*

fission, it is replicated and attached to the membrane at a site adjacent to the original copy. Continued growth of the cell membrane eventually separates the two copies of DNA, and a cell pinches in two daughter cells, each of which inherits a complete genome.

The genetic material (or genome) in a eukaryotic cell is organized into multiple chromosomes. For example, a human somatic cell has 46 chromosomes, and a reproductive cell (sperm/egg) has 23 (half of the somatic cell chromosomes). Each chromosome is organized in a long, linear

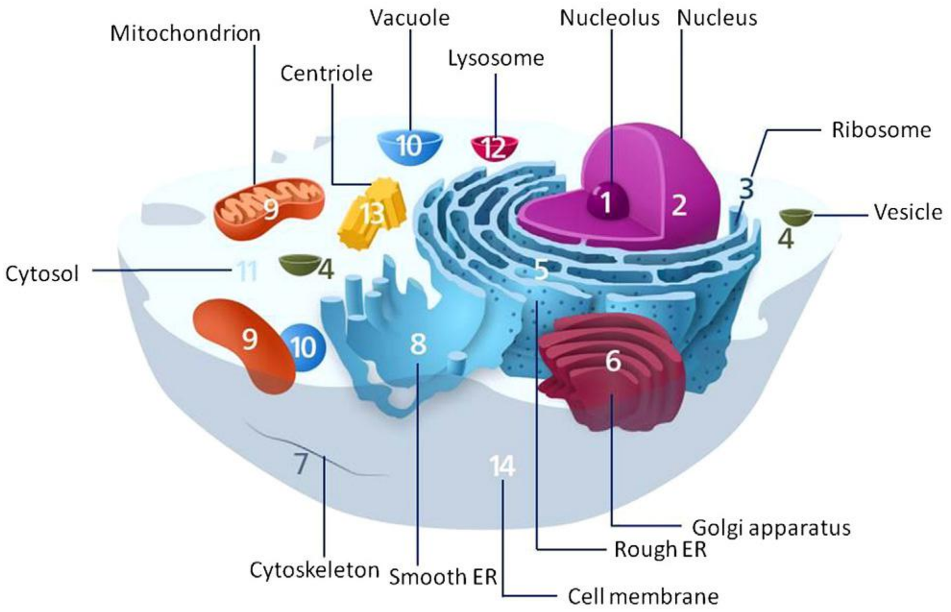


FIGURE 2.1 **The most common structures in an animal cell.** Genetic material, in the form of DNA, resides in the nucleus. Although vacuoles are included here, they generally exist in protists and plants. *Source: Reproduced from* Wikipedia Commons *(https://en.wikipedia.org/wiki/Organelle).*

DNA molecule representing a number of genes. Various proteins help to maintain the DNA structure. This DNA-protein complex is folded and coiled to produce the chromatid of a chromosome. The chromatid duplicates itself in the interphase of the cell cycle, and the two sister chromatids are held together by a specialized region called the centromere.

In a dividing cell, mitosis alternates with interphase, which takes up most of the time in each cell cycle. Interphase can be further divided into three periods of growth: the G_1 phase ("G" stands for gap, here the first gap), the S phase ("S" stands for synthesis), and the G_2 phase (the second gap). As it goes through interphase, the cell grows by synthesizing proteins and producing cytoplasmic organelles through all three subphases, while the chromosomes are duplicated only during the S phase. Once everything is prepared for cell division, the cell divides in the mitotic (or M) phase. How an animal cell completes its division during the M phase is described later.

Mitosis is conventionally separated into five stages: prophase, prometaphase, metaphase, anaphase, and telophase. During G_2 of interphase (late interphase), the nucleus, which contains one or more nucleoli, is enclosed in the nuclear envelope. Two centrosomes are present just outside of the nucleus. The replicated chromosomes (during the S phase) are still not distinguishable as loosely packed chromatin fibers. During prophase, the nucleoli disappear, and the chromatin fibers are condensed into discrete chromosomes, visible under a light microscope. Each duplicated chromosome consists of two sister chromatids held together by the centromere. As the two centrosomes move away from each other, early mitosis spindle (made of microtubules) begin to appear in the cytoplasm.

During prometaphase, the nuclear envelope is broken down, allowing the microtubule spindles to enter the nucleus and interact with the chromosomes. Some of the microtubules attach to a specific site, called the kinetochore, in the centromere region, preparing for chromosome movement. During metaphase, the chromosome's centromeres line up at the equator of the spindle, with each sister chromatid facing the spindle's opposite poles. During anaphase, the sister chromatids are separated from each other, and each one begins to move along the microtubules toward opposite poles. By the end of this phase, the two poles have identical collections of chromosomes.

At telophase, the nucleoli reappear and nuclear envelopes are reformed at the two ends of the cell, where the two sets of chromosomes have gathered. Mitosis, the division of the nucleus, is now complete and is followed immediately by cytokinesis, the division of the cytoplasm. Thereafter, two daughter cells separate shortly after the end of mitosis. In animal cells, cytokinesis starts with a cleavage furrow. Figure 2.2 describes these stages in an animal cell.

The timing and rate of cell division in different parts of an organism are deciding factors in normal growth, development, and maintenance. For example, human skin cells divide frequently, while muscle and nerve cells do not divide at all. Researchers have used cell culture to study the cell cycle and the factors that may influence it, such as nutrients, growth factors, and cell density.

Before cell division, the cell has to pass a checkpoint near the end of the G_1 phase. Cell size—more precisely, the cell's cytoplasmic volume-to-genome ratio—must reach a certain level to pass this checkpoint before DNA synthesis occurs at the S phase. Other environmental and

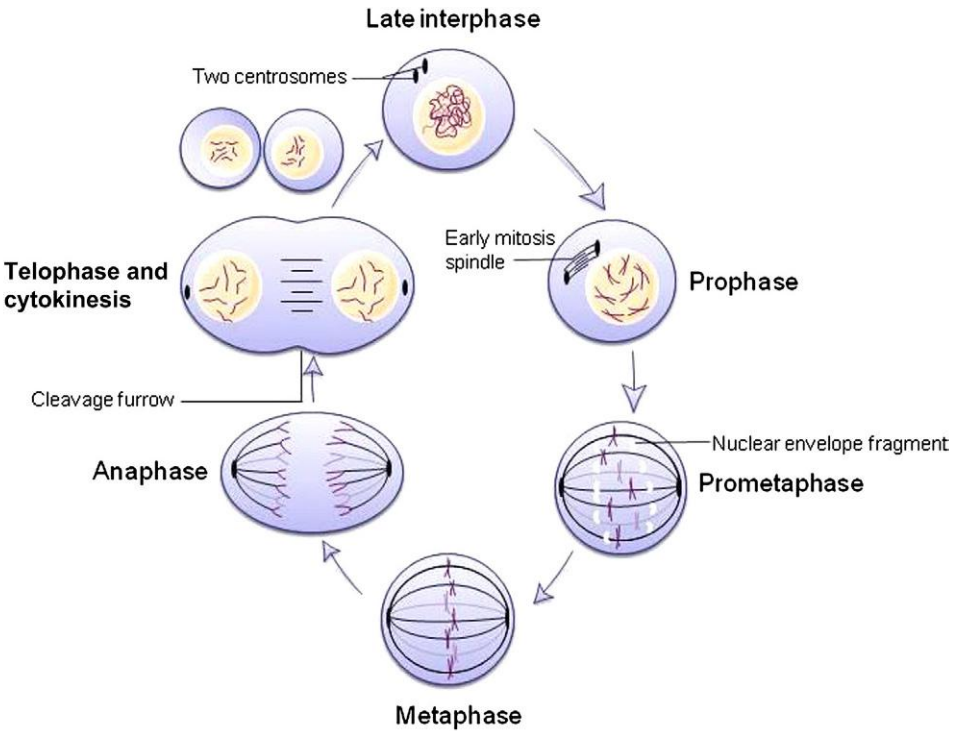


FIGURE 2.2 **Cell division—mitosis in an animal cell.** In this example, only seven chromosomes are shown.

developmental conditions must also be fulfilled to eventually initiate cell division. Protein kinases, which can be activated by cyclins (a class of regulatory proteins whose concentrations fluctuate cyclically during the cell cycle), regulate other proteins by phosphorylation; they also control the sequential steps of the cell cycle. In short, after the cell passes the checkpoint and is committed to reproduction, cell division is carried out as instructed by regulatory proteins (e.g., kinases, cyclins) in a cyclical manner.

2.3 MEIOSIS

Another form of cell division is meiosis, which occurs only in the ovaries and testes. Meiosis produces the reproductive cells, or gametes (i.e., ova and sperm cells), and reduces the chromosome number by half. Among the 46 human chromosomes, there are two of each type having the same length, centromere position, and band pattern. Such chromosomes are paired up and are referred to as homologous chromosomes, carrying genes controlling the same inherited traits. Each homologous chromosome in each pair is inherited from each parent. The 46 chromosomes can be organized into two sets of 23 chromosomes, with one set from the mother and the other set from the farther.

There are two distinct human sex chromosomes: X and Y. Females have a homologous pair of X chromosomes, while males have one X and one Y. X and Y are called sex chromosomes, and the others (i.e., 22 pairs) are called autosomes. Each gamete with a single set of 23 chromosome is a haploid cell. Fertilization of the gametes (a haploid sperm cell fuses with a haploid ovum, resulting in a fertilized egg known as a zygote) brings the chromosomal number back to 46. Somatic cells are derived from the zygote by mitosis during the development of a sexually mature adult. Diploid cells include the zygote and the somatic cells that have two sets of chromosomes.

Fertilization and meiosis are two features of sexual reproduction and are common to all organisms that reproduce sexually, although the details of sexual life cycles vary across species. The stages of meiotic cell division, using an animal cell as an example, are further illustrated next.

Interphase, during which each of the chromosomes replicates, is similar to mitosis. After chromosome replication, four daughter cells with only half as many chromosomes as the parent are formed through two meiotic divisions, meiosis I and meiosis II. During meiotic prophase I, two side-by-side homologous chromosomes pair up with each other in a process called synapsis. Chromatids of homologous chromosomes are criss-crossed in a process known as chiasma (see Figure 2.3). From one side this structure helps hold homologous chromosomes together; from the other side, it may result in significant genetic variations in the offspring (discussed in a later section). The two centrosomes move away from each other, and early spindle microtubules are formed between them that are similar to those observed during mitosis.

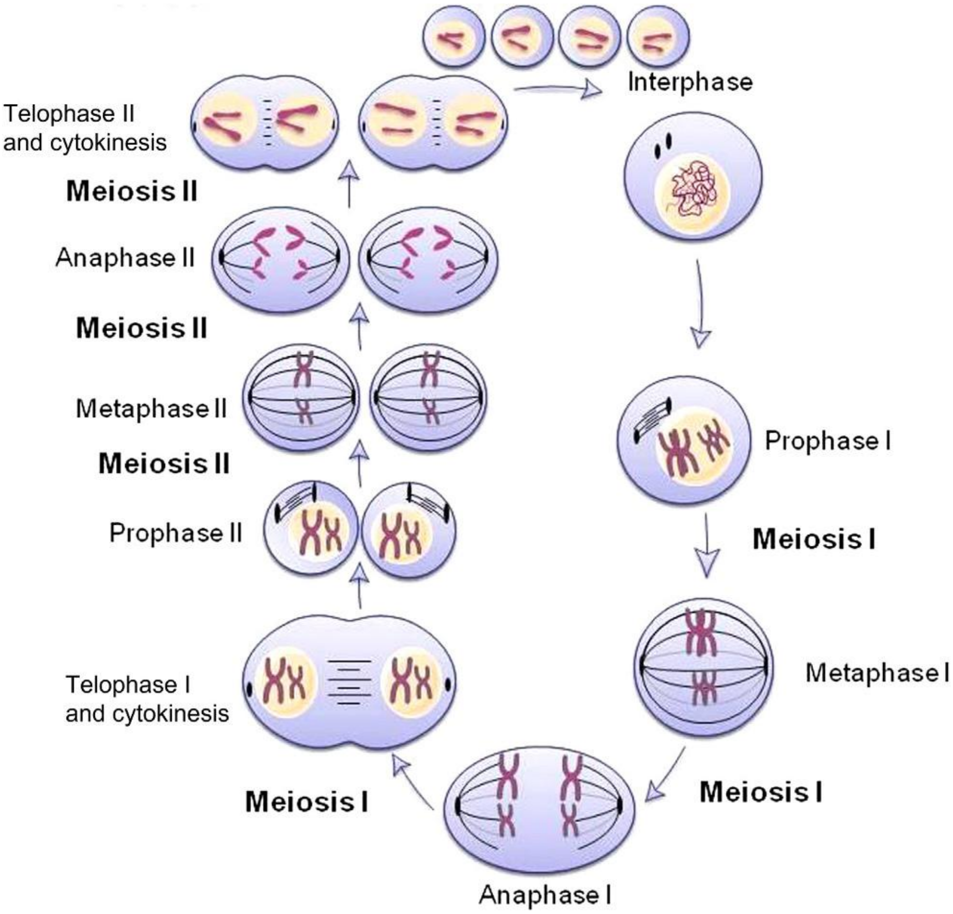


FIGURE 2.3 **Cell division—meiosis in an animal cell with a diploid number of 4 (or two sets of chromosomes).**

During metaphase I, the paired homologous chromosomes are arranged on the spindle equator. Spindle fibers from opposite poles attach to one chromosome of each pair, respectively. During anaphase I, homologous chromosomes move toward the two spindle poles; however, sister chromatids of each chromosome remain attached at the centromere and move together toward the same pole. During telophase I, the homologous chromosome pairs continue to be separated by the spindle. The homologues are finally split, and each pole now has a haploid chromosome set. Cytokinesis occurs at the same time and leads to the formation of two daughter cells at the end of meiosis I.

There is no further chromosomal replication prior to meiosis II, which also includes four subphases like the ones in meiosis I—prophase II, metaphase II, anaphase II, telophase II, and cytokinesis (see Figure 2.3). The differences are that, during metaphase II, chromosomes align in a similar manner to that observed in mitosis, with the spindle fibers attaching to the kinetochores of each chromosome's sister chromatids; during anaphase II, these sister chromatids move toward opposite spindle poles. Four daughter cells are formed following telophase II and cytokinesis, and each daughter cell has a haploid number of chromosomes compared to the zygote or the somatic cells. Table 2.3 compares mitosis and meiosis in an animal cell.

2.3.1 Genetic Variation Introduced by Meiosis: The Mechanisms

At the chromosomal level, a crossing over of two nonsister chromatids of homologous chromosomes during meiosis I (i.e., prophase I) results in individual chromosomes inheriting a portion of DNA from both mother and father. Such an exchange, or crossover, accounts for the recombination of linked genes on the same chromosomes. Independent assortment of chromosomes and random fertilization are also sources of genetic variability.

During meiosis I (i.e., metaphase I), a pair of homologous chromosomes orient independently of any other pair in terms of orientation relative to the two poles of the cell. Starting with two pairs of homologous chromosomes, there are four different combinations possible for a gamete. Based on the haploid number (n), the total number of combinations for the gametes generated by meiosis can be

TABLE 2.3 Comparison of Mitosis and Meiosis in an Animal Cell

	Mitosis	Meiosis
DNA replication	Occurs during interphase	Occurs during interphase
Division number	One division, including prophase, metaphase, anaphase, and telophase	Two divisions (meiosis I and II), each consisting of prophase, metaphase, anaphase, and telophase; homologous chromosomes joining along their length are unique to meiosis
Division outcome	Results in two daughter cells (diploid) genetically identical to the mother cell	Results in four daughter cells (haploid); genetically nonidentical to the parent cell
Significance	Derived from the zygote and has a positive effect on cell growth for repairing tissues	Gamete production; compensates chromosomes' doubling during fertilization by reducing the number of chromosomes to half; produces genetic variability in gametes

calculated as 2^n. In the case of humans, the haploid number is 23; thus, the number of all possible combinations in the gametes is 2^{23}. This independent assortment of homologous chromosomes produces about 8 million chromosomal combinations for human gametes. It produces another form of genetic recombination among unlinked genes that locate on separate nonhomologous chromosomes.

Random fertilization adds another layer to the creation of genetic variations. For a zygote to form, a human ovum representing one out of 8 million possible chromosome combinations is fertilized by a sperm cell, whose chromosome composition presents 8 million different possibilities as well. Therefore, without taking crossover into account, 64 (8×8) million diploid combinations for a zygote are possible.

The three mechanisms just described reorder the genes carried by individuals in a population; nevertheless, the fundamental factor that explains genetic diversity in a population is mutation.

2.4 MENDELIAN LAWS

Around 1857, Gregor Mendel began breeding garden peas in an abbey garden to study the mechanism of inheritance. In this section, we examine in detail the heredity principles developed by Mendel and the application of his model to the inheritance of human variations.

Mendel's choice of the pea as his experimental organism may be explained as follows. First, there are many varieties of peas. For example, flower color, seed color, and shape vary among different pea varieties, and such characteristics can all be used to observe inheritance patterns. Second, working with peas allowed Mendel to strictly control hybridization. Since carpels and stamens are almost completely enclosed by the petals of pea flowers, instead of allowing self-fertilization from the same plant, Mendel could cross-pollinate between different varieties (e.g., pea plants with different flower color) by removing the immature stamens and transferring pollen from another plant onto the emasculated flowers. In this way he controlled the parentage of the hybridized seeds.

In his breeding experiments, Mendel cross-pollinated purple-flower and white-flower pea plants and ensured that both plant types (either purple flowers or white flowers) were self-pollinated in the first place and that their offspring were of the same variety. This is known as true breeding (e.g., a purple-flower plant always produces offspring having purple flowers through self-pollination). Hybridization is defined as the mating of two true-breeding varieties. The true-breeding parents are referred to as the P generation, and the first generation of offspring is referred to as the F_1 generation. Self-pollination of the F_1 generation produces the F_2 generation. Mendel followed these generations in his experiments. It was after the quantitative analysis of the F_2 generation that two main principles of heredity were revealed: the law of segregation and the law of independent assortment.

2.4.1 The Law of Segregation

In his flower color breeding experiments, Mendel used the true-breeding peas, with purple and white flowers respectively, as the parental (or P) generation. Like their parents, the F_1 offspring all had purple flowers. Mendel allowed them to self-pollinate, and the white-flower peas reappeared in the F_2 generation. More specifically, Mendel observed 705 plants with purple flowers and 224 with white flowers. An approximate ratio for this observation is 3 (purple) to 1 (white). Along with flower color, Mendel examined the pattern of inheritance for flower position, seed color, seed

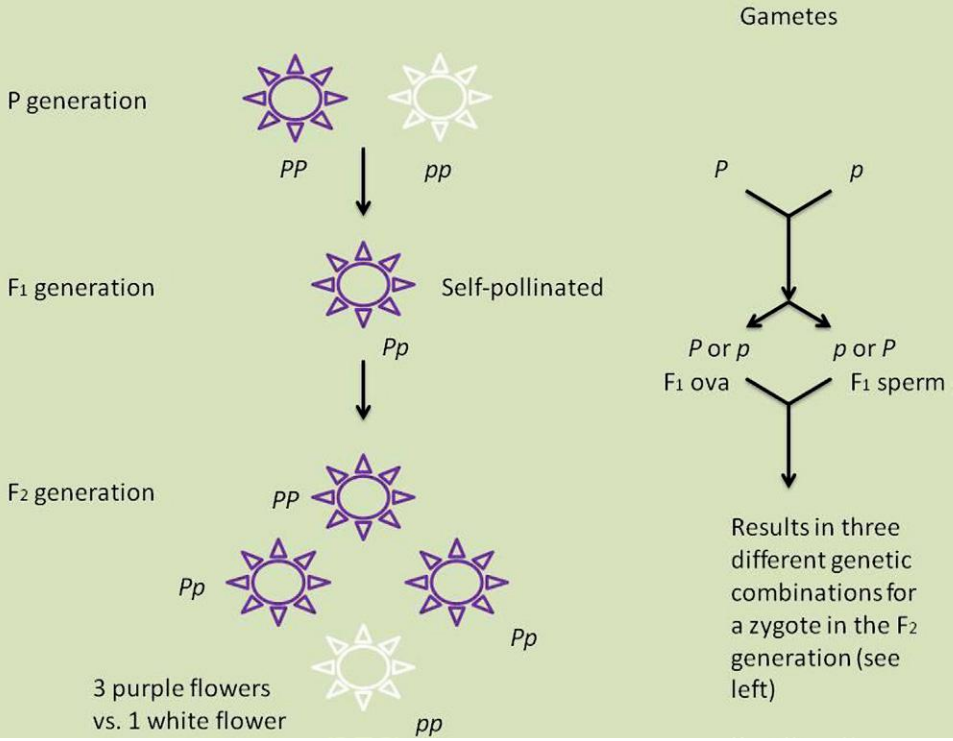

Gametes
P generation
PP
pp
P
p
F1 generation
Self-pollinated
Pp
P or p
p or P
F1 ova
F1 sperm
F2 generation
PP
Pp
Pp
Results in three different genetic combinations for a zygote in the F2 generation (see left)
3 purple flowers vs. 1 white flower
pp

FIGURE 2.4 **Mendel's law of segregation.** Shown is Mendel's crossing experiment for a single character (e.g., flower color).

shape, pod shape, pod color, and stem length; he observed that all of the pea plants followed the same pattern of inheritance as that seen in flower color, with a typical ratio of 3:1 in the F_2 generation.

Mendel developed a theory to explain the observed inheritance pattern in flower color that can be separated into four hypotheses:

- Different versions of genes account for variations appearing in the F_2 generation. For example, an alternative gene version is present in white flowers relative to the version in purple flowers. Different gene versions are now known as alleles.
- Two alleles from each parent are inherited to represent a trait. This is not difficult to understand given what we encountered during the discussion of meiosis: in a diploid cell, one homologous chromosome of a pair is inherited from each parent; a genetic locus is thus represented twice along the paired homologous chromosomes. The alleles may be matched at a homologous locus—for example, the true-breeding P generation in flower color—and the alleles may be different at the same locus—for example, the F_1 generation (see Figure 2.4).
- Two different alleles can be either dominant (fully expressed to control for an organism's appearance) or recessive (no noticeable effect on appearance). Thus Mendel's F_1 plants inherited a purple-flower allele from one parent and a white-flower allele from the other parent. That the F_1 plants are purple is possibly because the purple-flower allele (the *P* allele, as shown in Figure 2.4) has a dominant effect on flower color, while the white-flower allele (the *p* allele, as also shown in Figure 2.4) is a recessive allele and its effect on flower color is masked by the dominant *P* allele.
- The two alleles segregate during the formation of a gamete, which means that the reproductive cell (sperm or ovum) has only one of them. As in the F_1 plants, half of the gametes receive the dominant allele and half receive the recessive allele (Figure 2.4). The separation of alleles into gametes is actually the central idea of Mendel's law of segregation.

Figure 2.4 shows how the model of segregation explains the 3:1 ratio observed in the F_2 generation. Two classes of F_1 gametes are formed. When alleles segregate, half of the gametes receive a *P* allele (purple flower) and the rest receive a *p* allele (white flower). During self-pollination, the two classes of an ovum's gametes unite randomly with the two classes of a sperm's gametes, resulting in a total of four possible combinations in a fertilized egg. In the F_2 generation, one-fourth of the plants apparently have two *P* alleles, leading to the expression of purple flowers. Half of the plants have one *P* and one *p* allele and have purple flowers according to the third hypothesis. One-fourth of the F_2 offspring have white flowers with the two *p* alleles, expressing the recessive white flower feature. This accounts for the exact ratio of 3:1 observed in the F_2 generation.

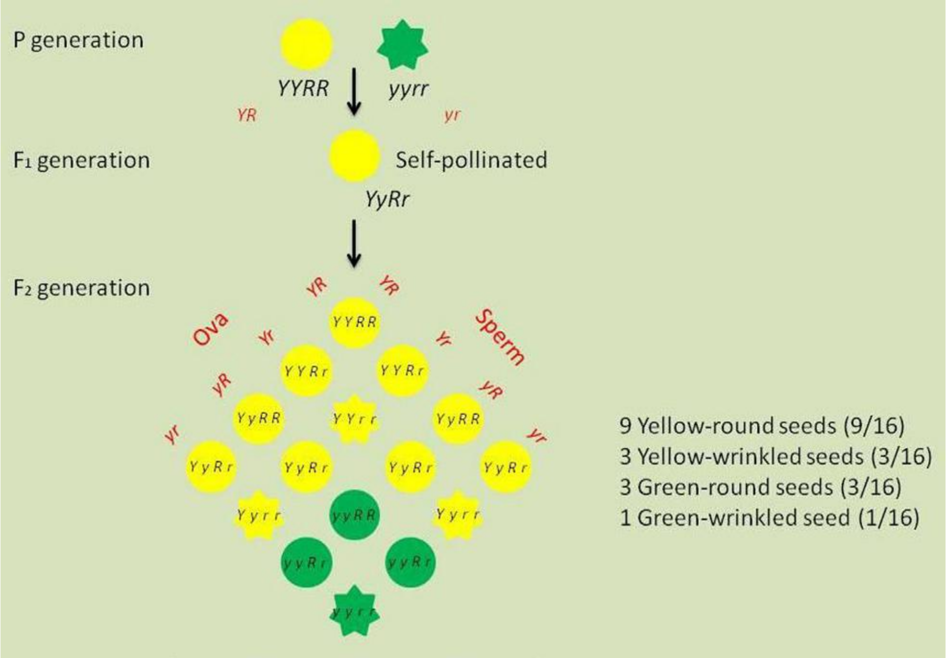


FIGURE 2.5 **Mendel's law of independent assortment.** Shown is Mendel's crossing experiment between true-breeding parents differing in two characters, seed color and seed shape.

A pair of identical alleles of a gene accounting for an inherited trait is said to be homozygous for that gene. Two different alleles at the same genetic locus representing a trait are said to be heterozygous at that locus. In the example of flower color in peas, *PP* and *pp* are homozygous for purple and white flowers, respectively, with *Pp* as the heterozygous type. Another concept that describes the observed physical appearance is the phenotype, whose genetic makeup is called its genotype. In the same case of flower color, both *PP* and *Pp* express the dominate phenotype of purple flowers, but they are different genotypes.

2.4.2 The Law of Independent Assortment

Mendel discovered the model of allele segregation into gametes using parental varieties that differ in a single character (e.g., flower color). He also conducted mating experiments on parental varieties differing in two characters (e.g., seed color and shape). From previous crossing experiments in single characters, including seed color and shape, Mendel already knew that yellow is a dominant trait (for the *Y* allele) relative to green (for the *y* allele), and that round is a dominant trait (for the *R* allele) relative to wrinkled (for the *r* allele). Breeding a plant with yellow-round seeds (*YYRR*) to a plant with green-wrinkled seeds (*yyrr*), produces F_1 offspring with yellow-round seeds (*YyRr*). Following self-fertilization of the F_1 plants, will the *Y* and *R* alleles still be transmitted together to the F_2 generation, or will the two allele pairs segregate independently of each other? To answer this question, we now take a look at the overall design and theoretical outcome of Mendel's experiment on crossing peas differing in both seed color and shape (Figure 2.5).

The experiment produces a ratio of 9:3:3:1 in the F_2 generation: nine yellow-round seeds, three yellow-wrinkled seeds, three green-round seeds, and one green-wrinkled seed (Figure 2.5). In this example, four gametes are formed given self-pollination of the F_1 generation: *YR*, *Yr*, *yR*, and *yr*. When these four ovum classes are fused with four sperm classes, 16 different combinations of gametes' alleles are generated in the F_2 offspring (Figure 2.5). The results of this experiment support the independent segregation of each pair of alleles into gametes. Indeed, Mendel obtained 315 yellow-round, 108 green-round, 101 yellow-wrinkled, and 32 green-wrinkled seeds from his experiment—a ratio of roughly 9:3:3:1. This pattern of inheritance is now defined as Mendel's law of independent assortment.

2.4.3 Mendelian Patterns of Inheritance in Humans

In modern genetics, Mendelian principles have been extended to more complex inherited traits than Mendel described. His peas have a relatively simple genetic basis—each character is determined by only one gene with two versions (or alleles), of which one is completely dominant. In fact, the relationship between genotype and phenotype for the majority of heritable traits is very complicated, but the laws of segregation and independent assortment can still be applied.

Recall from previous sections that pea characters, such as flower color, seed color, and seed shape, are completely controlled by their respective dominant alleles: *P*, *Y*, and *R*. In the case of flower color, Mendel's heterozygous F_1 offspring (*Pp*) all had purple flowers because the *P* allele shows a completely dominant effect over the recessive *p* allele. However, some heterozygous genotypes may cause incomplete dominance, meaning that the appearance of those heterozygous individuals has an intermediate phenotype between the phenotypes of the parent generation. Along with complete and incomplete dominance patterns of inheritance, there is co-dominance, in which a heterozygous individual expresses both phenotypes of the two alleles. Furthermore, the dominant effect of one allele on a phenotype is reflected by the mechanisms/pathways from genotype to phenotype, which does not imply the ability of one allele to mute another at the DNA level or the abundance of that allele in a population.

The relationship between genotype and phenotype may also be explained by multiple gene alleles (instead of only the two reported by Mendel); pleiotropy (i.e., one gene affects several phenotypic traits); epistasis (i.e., a gene at one locus interferes with the expression of a second gene at a different locus); and environment (e.g., nutrition affects human height and sun exposure affects skin color). Some characters, such as human height and skin color, result from an additive effect of two or more genes in a continuous fashion. These characters are said to be polygenic. Environment influences these polygenic traits, which are thus known as multifactorial.

Mendel's discovery of the patterns of individual gene transmission from parents to offspring led to Mendelian models, which were developed and broadly used to explain the inheritance patterns of epistasis and polygenic characters. In the case of human inheritance, family pedigree analysis shows evidence that supports Mendelian patterns, with some following those patterns recessively (i.e., inherited from the heterozygous genotype) or dominantly (i.e., inherited from the dominant allele). Besides these simple Mendelian disorders, humans are more prone to diseases that have a multifactorial basis. For example, heart disease, diabetes, cancer, and many others diseases are the result of multiple genes, interactions between genes, and interactions between genes and the environment. Using technologies developed in recent years, studies have been conducted on the genetic and environmental components of multifactorial human traits. These will be discussed in later chapters.

2.5 PUBLIC DATABASES FOR BIOMEDICAL RESEARCH IN HUMANS

Groups of researchers have formed large-scale, international collaborations with the goal of creating accessible public databases of human genetic or epigenomic data. These databases will aid basic biomedical research that promotes understanding of the biology of human health and disease. Such projects include the International HapMap Project, the 1000 Genomes Project, the Encyclopedia of DNA Elements (ENCODE), and the Roadmap Epigenomics Project.

The International HapMap Project (http://hapmap.ncbi.nlm.nih.gov/) focuses on the identification of the common patterns (haplotypes) of genetic variations and the determination of tag single-nucleotide polymorphisms (SNPs) that represent them. Analysis of 1.6 million SNPs in 11 populations of African, European, and Asian ancestries has been carried out, with a total sample size of 1184 (HapMap Phase I, II, and III, called "HapMap 3") [1]. This integrated dataset effectively captures common variants with minor allele frequency (MAF) of $\geq 5\%$ in the related populations, and it has improved imputation accuracy for variants with lower frequencies ($\text{MAF} \leq 5\%$). The project provides a robust reference panel for studying human genetic variation across various diseases and other common complex traits [1].

In October 2012, the 1000 Genomes Project (www.1000genomes.org/) Consortium announced the sequencing data from 1092 human genomes for 14 populations drawn from Europe, East Asia, Sub-Saharan Africa, and the Americas. This includes low-coverage whole-genome sequence data (2–6 times), targeted deep-exome sequence data (50 times), and dense SNP genotype data. Ultimately, ~2500 individuals from 26 populations will be sequenced. The aim of the 1000 Genomes Project is to generate a more comprehensive catalog of genetic variations by capturing SNPs at a frequency of 1% in the human genome [2].

ENCODE (http://genome.ucsc.edu/ENCODE/), funded by the National Human Genome Research Institute (NHGRI), began in September 2003 with the aim of identifying all functional elements in the human genome. These include protein-coding regions, RNA-transcribed regions, transcription factor binding sites, chromatin structure and histone modification sites, and DNA methylation sites. So far, the overall results reflect only a small portion of the potential functional elements encoded in the genome [3]. Other factors, modifications, and cell types can be studied and added to create a larger and more comprehensive database for understanding the functions of the human genome. Also, integration with other, related projects (for example, the Roadmap Epigenomics Project) will allow a deeper analysis of genes and regulatory factors and an understanding of the mechanisms of regulation [3].

The Roadmap Epigenomics Project (www.roadmapepigenomics.org/) was developed for documenting human epigenomic data to map DNA methylation, histone modification, chromatin accessibility, and small RNA transcripts using next-generation sequencing technologies in stem cells and *ex vivo* normal tissues. These are representative of the tissues involved in human disease. Normal epigenomes are provided and serve as a reference for possible combined investigation with other projects. The raw sequence data, epigenomic signatures, and integrated maps are released rapidly into the public domain. Standardized protocols and analytical tools are also provided as guidelines to the scientific community.

2.6 CONCLUSION

In this chapter, we first looked at the cell structure and function of organelles using an animal cell. We then moved to the topics of cell division, including mitosis, which produces two daughter cells identical to the parent cell, and meiosis, which generates four daughter cells different from the parent cell as well as from each other. Last, we reviewed Mendel's crossing experiments on garden peas and the fundamental concepts of inheritance revealed by his experiments. The impact of these principles on human genetics was also reviewed, followed by information on several publicly available databases for human biomedical research.

GLOSSARY

Allele one version of a gene at a given locus along a chromosome

Band During cell division, the specific staining pattern (light and dark stripes) is observed for chromosomes. Such patterns can help in identifying chromosomes and evaluating their structure

Binary fission a type of cell division that applies to prokaryotes, such as bacteria. One cell is divided in two, with two complete and identical genomes present in the two daughter cells

Centromere a constricted chromosomal region that holds sister chromatids together during cell division. The centromere also separates a chromosome into a short arm and a long arm

Character a heritable feature that varies among individuals

Chromatid during the S phase of the cell cycle, each chromosome is duplicated. Following replication, each chromosome consists of two identical sister chromatids, which contain matched copies of the DNA sequence at every locus

Chromatin fiber the 30 nm coil of DNA and proteins that forms the basic structure of chromatin

Cytokinesis the division of the cytoplasm. Cytokinesis immediately follows the division of the nucleus

Cytoskeleton protein filaments within cells that maintain cell shape and cell movement

DNA the molecule that encodes the genes responsible for the structure and function of an organism. Genetic information carried by DNA molecules is transmitted from one generation to the next

DNA methylation methyl groups, consisting of one carbon atom and three hydrogen atoms, attach to DNA molecules at cytosine bases, resulting in reduced transcription of the gene

Endomembrane system all membranes, which are suspended in the cytoplasm within a cell, can divide the cell into structural and functional compartments/organelles. In a eukaryotic cell, endomembrane organelles include the nuclear envelope, the endoplasmic reticulum, the Golgi apparatus, lysosomes, vacuoles, vesicles, and the cell membrane. The endomembrane system does not contain mitochondrial membranes

Epigenomics the study of global gene expression changes caused by epigenetic processes that do not involve any change in the underlying DNA sequence

Epistasis the expression of one gene depends on the expression of one or more independently inherited genes

Eukaryote an organism whose cells contain a nucleus and other membrane-bound components. Many unicellular organisms, such as protozoa, are eukaryotes. All multicellular organisms—animals, plants, and fungi—are eukaryotes

Exome the part of the genome that corresponds to exons, which are the coding sequences of DNA present in mature messenger RNA. The exome constitutes about 1% of the total human genome

Gamete a cell/gamete that carries half the genetic information of an individual fuses with another cell/gamete in organisms that reproduce sexually. There are two types of gametes: a sperm is produced by a male and an ovum is produced by a female

Gene the basic physical unit of heredity. For example, a segment of DNA that is arranged in a linear manner along a chromosome is a gene. A gene codes for a specific protein that has a particular characteristic or function

Genome the entire genetic material found in a cell. In the case of humans, the genome includes 23 pairs of chromosomes in the nucleus and the chromosomes contained in a cell's mitochondria

Genotype the genetic make-up of a cell or organism. Also, a set of alleles inherited at a locus

Haplotype a set of alleles or single-nucleotide polymorphisms found on the same chromosome, tending to be inherited together

Imputation in statistics, imputation is a process of replacing missing data with other estimated probable values. In genetics, imputation is a method used to fill in missing genotypes in a study dataset

Kinetochore a protein structure on chromatids, allowing the spindle fibers to attach and pull the sister chromatids apart during cell division

Linked genes genes on the same chromosome tend to be segregated and inherited together

Locus the location or the physical site of an individual gene or DNA sequence on a chromosome

Minor allele frequency the frequency of the least common allele in a given population

Mutation an alteration in a DNA sequence from its natural state. The effect of any mutation might be deleterious or benign

Next-generation sequencing a high-throughput sequencing technology relative to capillary electrophoresis–based Sanger sequencing, producing thousands or millions of sequences simultaneously. This technology can be applied to DNA analysis, RNA analysis, and gene regulation analysis

Organelle a subcellular structure within a cell. An organelle is usually enclosed by its own membrane for executing its specific function independently of other organelles in a cell, particularly a eukaryotic cell

Organism any living system that has the ability to grow, reproduce, maintain homeostasis within itself, and function independently. An organism may be either unicellular or multicellular. All organisms on Earth can be further divided into eukaryotes and prokaryotes

Phenotype a gene's expression of observable physical and/or biochemical characteristics. Some phenotypes are heavily influenced not only by genes but also by the environment

Pleiotropy multiple physical effects caused by a single gene. The underlying mechanism of pleiotropy may be a gene involved in different metabolic pathways that contribute to different phenotypes

Prokaryote an organism lacking a membrane-enclosed nucleus. Most prokaryotes are unicellular. No membrane-bound organelles are present in the prokaryotic cells. The intracellular components of a prokaryotic cell are gathered together and enclosed only by the cell membrane

Recombination the new combinations of the traits in offspring that are inherited from its parents, now more commonly known as the exchange of a DNA segment between two homologous chromosomes during meiosis, which causes new combinations of hereditary material in a gamete

RNA a molecule synthesized from the DNA template, with the alternating sugar ribose to replace the deoxyribose present in a DNA molecule. Different types of RNA exist in a cell: messenger RNA, ribosomal RNA, transfer RNA, and some small RNAs

Sequencing the technique for determining the exact DNA sequence or nucleotides (A, C, G, and T). DNA-sequencing technology is developing fast and becoming less expensive. For example, whole-genome and whole-exome sequencing is now available for studying the association between a DNA sequence and a disease

Single nucleotide polymorphism a change in a single nucleotide at a genetic locus with a minor allele frequency of 1% or higher in a population

Somatic cell the cell forms the body of an organism. For example, in humans, somatic cells form all internal organs, skin, bones, and connective tissues. Somatic cells are diploid

Synapsis the pairing of two homologous chromosomes at the early stage of meiosis. This allows the matching of homologous chromosomal pairs and prepares for possible crossover between them

Trait a variant of a character, which may be determined by genes, the environment, and the interactions between the two

ACKNOWLEDGMENTS

The information in Campbell's book and on the publisher's website were most useful while preparing this chapter: Campbell N.A. Biology (4th ed.), Benjamin/Cummings, 1996. The website www.nature.com/scitable is provided by the Nature Publishing Group.

REFERENCES

[1] Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. Nature 2010;467(7311):52–8.

[2] Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491(7422):56–65.

[3] Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489(7414):57–74.

Chapter 3

The Human Genome, Gene Regulation, and Genomic Variation

Kathleen M. Murphy, Aislinn Cooper, and Edward S. Tobias
Medical Genetics, School of Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Chapter Outline

3.1 Overview: Structure of the Human Genome 41
3.1.1 Chromatin Structure 42
3.1.2 The Nucleosome 43
3.2 Gene Structures 44
3.2.1 Number of Genes 44
3.3 Gene Expression 45
3.3.1 Transcription 45
3.3.2 Translation 45
3.4 Gene Regulation 46
3.4.1 Gene Regulation by miRNAs 47
3.4.2 mRNA Targets 48
3.4.3 Chromatin Modifications and Gene Expression 50
3.4.4 Nucleosome Positioning and Occupancy 50
3.5 Genomic Variation 51
3.5.1 Single-Nucleotide Polymorphisms 53
3.5.2 Copy Number Variants 54
3.5.3 The ENCODE Project and the Characterization of Human Variants 54
3.6 Conclusion 56
References 56

3.1 OVERVIEW: STRUCTURE OF THE HUMAN GENOME

The size of the human genome in the chromosomes of the nucleus is approximately 3100 million (3.1×10^9) base pairs (bp). This is the total length of the DNA in a haploid (i.e., half) set of human chromosomes, and it is based on the GRCh37-computed genome assembly (as updated in June 2013). Only around 1–1.5% of the human nuclear genome constitutes coding DNA or protein-encoding genes. This is consistent with the fact that there are approximately 23,411 protein-encoding genes (including alternative sequences, based on the GRCh37 dataset) and that the size of the coding region of a nuclear human gene is usually 1–3.5 kbp.

At the time of writing, there are also reported to be 9719 short noncoding genes and 14,335 long noncoding genes (although the Ensembl Genome Browser should be checked for updated figures; see [1,2]). Approximately 20% of the genome consists of introns and untranslated regions of genes, as well as other sequences that are related to genes, such as pseudogenes. Most of the remainder of the genome, constituting 75% of its total size, is extragenic, however, and as much as around 55% of the genome comprises repeated sequences. This repetitive sequence is predominantly composed of transposons (transposable elements) that can insert multiple copies of themselves in the genome in scattered locations. Interestingly, the proportion of the human genome that consists of repetitive sequences (i.e., 55%) is rather higher than in other organisms, such as the worm and the fly, in which the corresponding proportions are 7% and 3%, respectively. However, through rearrangement of regions of the genome by recombination, these repetitive sequences have permitted the structural modification of existing genes and even the creation of new ones.

Nucleic acid is the carrier of genetic information, and it is of two principal types: DNA (deoxyribonucleic acid) and RNA (ribonucleic acid). Both consist of a sugar–phosphate

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00003-7
© 2014 Elsevier Inc. All rights reserved.

backbone together with projecting nitrogenous bases. In DNA, the sugar is deoxyribose, whereas in RNA it is ribose. The bases are attached to the 1′ (one prime) position of each sugar, and the backbone is formed by the linking of 3′ and 5′ hydroxyl groups by a phosphate. A nucleotide is the unit consisting of a purine or pyrimidine base together with the attached sugar and phosphate group(s). In DNA, there are two pyrimidine bases, thymine (T) and cytosine (C), and two purine bases, adenine (A) and guanine (G). RNA also contains the bases cytosine, adenine, and guanine, but it contains uracil (U) instead of thymine (see [3]).

A molecule of DNA is composed of two strands, or chains, of nucleotides, which are coiled clockwise around each another. They form a double helix, with the two chains running in opposite directions (i.e., 5′ to 3′ for one and 3′ to 5′ for the other). There are 10–10.5 nucleotides per complete turn. The strands are held together by hydrogen bonds between the bases—i.e., between A in one chain and T in the other or between G and C. This base pairing is very specific, and, because A:T and G:C pairing is obligatory, the parallel strands must be complementary to one another. Thus, if one strand reads 5′-ATGC-3′, the complementary strand must read 5′-GCAT-3′ (not 5′-TACG-3′). Wide variation exists in the (A + T): (G + C) ratio, with higher plants and animals tending to have an excess of A + T. In humans, for instance, the ratio is 1.4:1.

DNA's unit of length is the base pair (bp), with 1000 bp in a kilobase (kb) and 1,000,000 bp in a megabase (Mb). Its total length in a half (haploid) set of human chromosomes is approximately 3100 Mb (3.1×10^9 bp), as mentioned previously, and as the distance between base pairs in the DNA helix is 0.34 nm, the total length of haploid DNA if extended per cell is 1 m.

Each protein-coding gene is usually present as just one copy in the haploid genome. Well-known exceptions, however, are α-globin, histone, and ubiquitin genes. The average size of a gene is 53.6 kb with an average intergenic distance of 60–70 kb. The majority of the genome is noncoding and repetitive and has no proven function. Work conducted in 2012 by the ENCODE Project Consortium suggests that many noncoding regions may have biological activities, at least in one or more cell lines used in experiments. Repetitive DNA is subdivided into tandem (multiple adjacent) repeats (satellite DNA) and interspersed (single) repeats.

Tandem repeats are subdivided according to their length. Microsatellite repeat arrays are under 1 kb in length, and the most common repeat motifs are A, CA, AAAN (where N is any nucleotide), AAN, and AG. The arrays of CA motifs commonly have 10–60 repeats and are found approximately every 36 kb. The repeat number on corresponding chromosomes typically differs, and these common genetic differences, or polymorphisms, have often been used to track the inheritance of that region of each chromosome. Single-nucleotide polymorphisms (SNPs) can now be used for this purpose.

Minisatellite repeat arrays are usually 1–30 kb in length and have longer repeat motifs than do microsatellite repeats, but again they show marked variation in repeat number. Macrosatellite repeats are even longer and may be many megabases in length. They are found at the ends (telomeres) of the chromosomal arms and in the central chromosomal constriction (the centromere). Length variation is common and, in fact, accounts for visible differences in the size of chromosomal centromeric regions.

In contrast, interspersed repeats usually occur as single copies and are subdivided according to their length. Short interspersed repeats (SINES) are less than 500 bp in length, and the commonest of these is the Alu. SINE length is approximately 300 bp, and SINES contain a cutting site for the restriction enzyme *Alu*. Alu repeats are very common (constituting 11% of all DNA), occurring every 5–10 kb. Long interspersed repeats (LINES) vary from 500 bp to 10 kb (most often 6–8 kb) in length and include three families that in total constitute about 20% of the human genome. The most common is the LINE-1 (L1) element, which, like other LINES, possesses structural similarities to retroviruses. It can replicate itself and insert a copy elsewhere in the genome. This retrotransposition may disrupt a gene at the new location and cause a genetic disorder (e.g., hemophilia).

Single-copy and repeat DNA is double-stranded, whereas RNA is single-stranded and subdivided into several types. Transfer RNAs (tRNAs) are synthesized directly on a DNA template and, although 61 types might be expected, surprisingly only 49 are found because some tRNAs can bind to more than one mRNA codon. This results from relaxed base pairing ("wobble") at the third-base position of the codons. The DNA templates for tRNAs tend to occur as multiple copies, which may be clustered or dispersed. Ribosomal RNAs (rRNAs) are also synthesized directly on DNA templates, which again occur as multiple clustered copies. These are located in the nucleolus organizer regions on the short arms of chromosomes 13–15, 21, and 22, and also on chromosome 1. The ribosomal RNAs are synthesized as large precursors in the nucleolus and then enzymatically cleaved.

3.1.1 Chromatin Structure

DNA and proteins are assembled and folded into structures that allow for compaction of the entire genome into the microscopic cell nucleus. This arrangement of DNA wound around proteins and into further higher-order structures is known as chromatin. At the most basic level, classically known as the "beads on a string" structure, 146

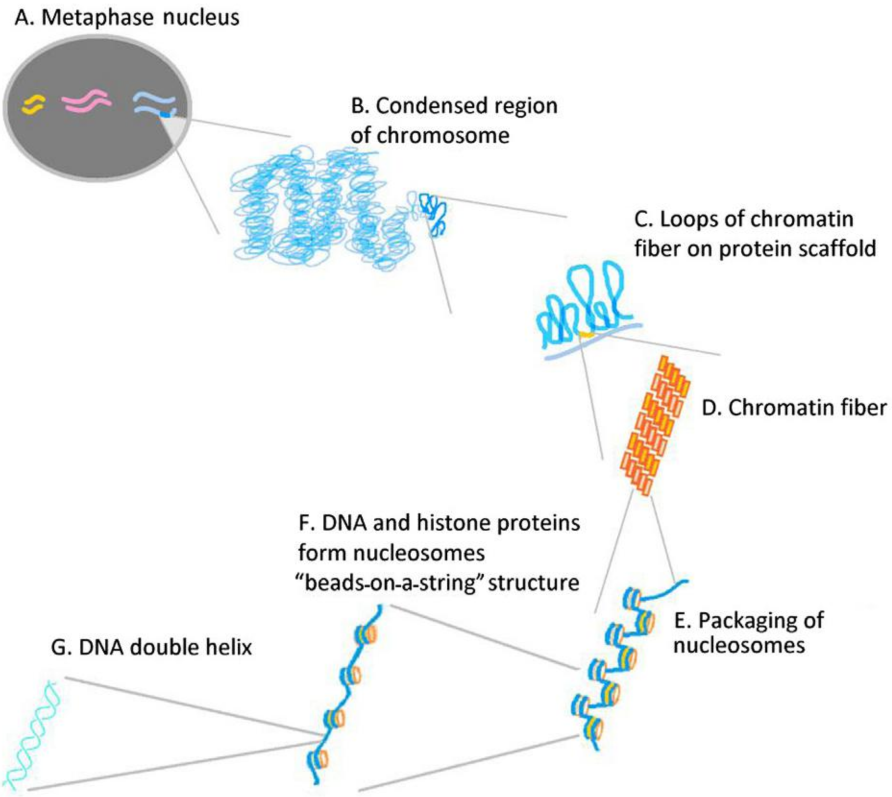

A. Metaphase nucleus
B. Condensed region of chromosome
C. Loops of chromatin fiber on protein scaffold
D. Chromatin fiber
F. DNA and histone proteins form nucleosomes "beads-on-a-string" structure
E. Packaging of nucleosomes
G. DNA double helix

FIGURE 3.1 **Organization of DNA within the nucleus.** Chromosomes are condensed in the metaphase nucleus and align along the metaphase plate in preparation for cell division. Looping fibers of chromatin on a protein scaffold are further folded into these condensed regions; they are made up of packaged nucleosome structures. Nucleosomes consist of 146 bp of DNA double-helix and histone proteins, appearing as "beads on a string" when viewed by electron microscopy. *Source:* [4].

base pairs of DNA are wound about 1.65 times around an octamer of histone proteins, constituting the nucleosome [5]. As illustrated in Figure 3.1, the organization of different levels of DNA into higher-order structures to achieve very high levels of compaction is complex (see [4,6]).

Chromatin structure is extremely dynamic and varies between loci on chromosomes and between tissue and cell types. To understand the manner and mechanism of this structural variation, first we must understand the different chromatin states that have been observed. Chromatin in loose or relaxed structures associated with active genes is commonly known as *euchromatin*, whereas tight or closed structures that tend to correspond to silent regions of the genome are referred to as *heterochromatin*. Heterochromatin can be subdivided into two categories: *constitutive* and *facultative* (see [4]). The first type is always found in a closed conformation and is therefore silent and not reversible; it will always exist in this closed state in an organism. Facultative heterochromatin, however, can be reversed to euchromatin and the genes in this region can therefore be expressed, often in response to signaling or external stimuli in processes such as development. Open regions are thought to be more actively expressed because of the ability of the transcription complexes and machinery to access the DNA and thus transcribe the genes in such a region. By this same reasoning, the genes in closed structures cannot be transcribed, as the interactions required between the DNA and other transcriptional components are not physically possible.

It is by these mechanisms that tissue-specific expression of particular regions of the genome is carried out. Regulation and control of the orchestration of such complex nuclear organization is mediated through chemical modifications of the histone tails in the nucleosome structure; these histone tails induce associations with further proteins in the cell and modify chromatin structure.

3.1.2 The Nucleosome

As previously mentioned, the nucleosome consists of DNA and an octamer of histone proteins; it is the basic component of chromatin structure. The histone proteins and their

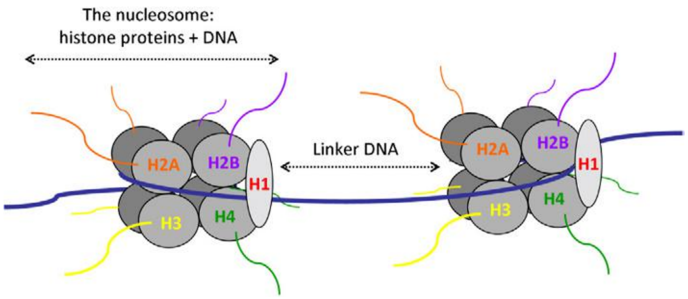


FIGURE 3.2 **The structure of the nucleosome.** The N-terminal tails of the histone proteins are not contained in the core of the DNA-histone complex; they project outward and thus can be easily accessed and modified by enzymes to ultimately alter overall chromatin structure in one particular genome region. Such modifications vary between regions in different cells and tissues. *Source:* [6].

modifications facilitate the opening and closing of chromatin structures, permitting or preventing gene transcription (see [6]). As seen in Figure 3.2, the N-terminal tails of the histone proteins protrude out of the nucleosome structure, allowing these residues to be subject to post-translational chemical modifications. Combinations of the modifications to specific residues on these tails, known as the "histone code," influence the DNA and chromatin interactions that can take place.

3.2 GENE STRUCTURES

Gene size is highly variable, ranging from small genes such as the globins to enormous genes such as titin and dystrophin (the largest, at over 2 million base pairs in length). A typical protein might contain about 300 amino acids, which are encoded by 900 bp of DNA. The vast majority of genes consist of protein-encoding exons alternating with noncoding segments, called intervening sequences or introns. The function of introns is still largely unknown, and they range from 50 bp to more than 10,000 bp in length. As a result of the presence of additional sequences, including these intervening sequences and also 5′ and 3′ flanking sequences, human genes are generally much larger than would be calculated from the encoded protein's amino acid sequence alone.

The initial mRNA (heterogeneous RNA or hnRNA) is a transcript of the gene that includes not only the exons but also the intervening sequences and some flanking sequences. The introns are removed from the hnRNA before it enters the cytoplasm by the process known as splicing. This requires the action of splicing proteins, which recognize particular sequences that are fairly well conserved at sites around the intron/exon junctions. Thus the sequence of an intron commences with GT (the 5′ splice donor site) and ends with AG (the 3′ splice acceptor site). Mutations in the bases adjacent to the exon/intron boundaries interfere with mRNA splicing and can lead to genetic disease by, for example, causing retention of an intron in the mature mRNA or abnormal removal of an exon. (see [3])

Each end of the mRNA is modified. At the 3′ end, the 3′ untranslated region (3′ UTR) extends from the protein translation stop codon to a poly(A) tail. This tail of 100–200 adenylate residues is not encoded in the DNA but is added enzymatically to aid cytoplasmic transport and, possibly, mRNA stability. At the 5′, a different modification is made: the end is protected or "capped," with a methyl guanylate residue, and the 5′ UTR extends downstream from this cap site to the beginning of the protein-encoding sequence.

Following ribosomal translation, many proteins are still not in their final form. Subsequent post-translational alterations include the formation of disulphide bonds (which can be structurally important for the protein), glycosylation, and proteolytic cleavage. Each step in the production of the final protein is important, as many proteins are highly dependent for functioning on their exact three-dimensional shape (or "tertiary structure"). This in turn is determined by the amino acid sequence and post-translational modifications of the protein. In general, acidic and basic amino acids are found on the surface of a folded protein, with hydrophobic amino acids being internally oriented. The effect of an amino acid substitution (and the degree of consequent disruption to the protein's tertiary structure) thus depends upon the amino acid's location relative to the active site of the protein. It can also depend, however, on the change in amino acid charge or the amino acid's affinity for water molecules.

3.2.1 Number of Genes

Chromosomes contain markedly different numbers of genes. For instance, the largest chromosome, chromosome 1, contains 2079 (the most protein-encoding genes), the X chromosome contains 830, and the Y chromosome contains 72 (the fewest protein-encoding genes). In general, genes are clustered in randomly distributed areas in the genome, with long regions of noncoding DNA in between.

Ancestral small duplications may explain this clustering. There might be subsequent divergence of duplicated genes' functions, and, in this process, some genes become nonfunctional copies, termed pseudogenes; some retain similar functions (e.g., the red-green color vision genes); and some develop novel functions as a result of small sequence changes or exon rearrangement.

The different (often computational) methods used to identify sequences as genes, and the subtypes of genes that are included, greatly affect the precise gene total. In addition to at least 23,411 protein-encoding genes, there are believed to be at least several thousand that code for RNA molecules but do not encode polypeptides. These RNA genes include at least 727 ribosomal RNA (rRNA) and 131 transfer RNA (tRNA) genes. A surprising number of other RNA genes are also now known to be present.

While their physiological roles are not yet as clearly understood as those of mRNAs, rRNAs, and tRNAs, a number of additional, intriguing RNA molecules are generally believed to be involved in the regulation of gene expression. These include, for instance, at least 2578 microRNA genes (described in detail later), 903 small cytoplasmic RNA (scRNA) genes, and 2019 small nuclear RNA genes (snRNAs). SnRNAs include RNAs that participate in splicing and a subclass of 1173 small nucleolar RNA (snoRNA) genes. Instead of being transcribed from separate genes, many snoRNAs are, interestingly, processed from introns spliced out from other genes. In contrast, scRNAs are usually found in complexes termed small cytoplasmic ribonucleoproteins (scRNPs), in association with cytoplasmic proteins. An example is the so-called signal recognition particle.

MicroRNAs (or miRNAs) are short single-strand RNA molecules of 21–23 nucleotides that regulate the expression of other genes. It is believed that regulation is mediated by the binding of the miRNAs to mRNAs especially in the 3′ UTR. This can cause the degradation of these mRNAs or the prevention of their normal translation into proteins. Although miRNAs do not encode proteins themselves, the large number of reports on the use of miRNA profiles (like mRNA profiles) as biomarkers of specific disease phenotypes should produce improved diagnostic and prognostic information to clinicians. (See Section 3.4 of this chapter.)

3.3 GENE EXPRESSION

3.3.1 Transcription

The process of transcription can be subdivided into three stages: *initiation*, *elongation*, and *termination*. Studies in the last five years have shown that these three stages coincide with the three steps involved in mRNA processing: 5′ capping, splicing, and 3′ processing.

Initiation of transcription is controlled by transcription factors (TF), which form *initiation complexes* with other TFs and proteins to repress or stimulate the transcription of a gene. TFs in the cell have multiple functions in the regulation of many different genes; thus their role is entirely dependent on the context in which they are found. TFs associate with particular sequences to either repress or stimulate gene expression—for example, the TF TATA-binding protein associates with the TATA box element located about 20–30 bp upstream of the transcription start site (TSS), bending the DNA double helix ready for subsequent assembly of the initiation complex. This complex then positions RNA polymerase II (RNAPII) at the TSS. The addition of helicase allows the unwinding of the DNA at the TSS so that transcription by RNAPII can commence.

The elongation step is a continuation of the primary initiation step in that this is when transcription progresses and a transcript is synthesized from the template strand. The C-terminal domain of RNAPII becomes phosphorylated as elongation occurs. Subsequent to initiation, a mechanism known as mRNA capping occurs, in which about 20–25 nucleotides of DNA are transcribed. The cap structure is added to the 5′ end of the growing mRNA via a short series of enzymatic reactions activated by the phosphorylated C-terminal domain of RNAPII. The cap has multiple functions: one is marking the TSS and another is the maturation of the mRNA and its stability in the translation step of protein synthesis. The general TFs that form the initiation complex dissociate.

Following this, elongation factors bind, forming what is known as the *elongation complex,* which facilitates the progression of RNAPII along the DNA. It is this complex that synthesizes the transcript. As elongation progresses, splicing factors are subsequently recruited. Co-transcriptional splicing of the nascent transcript occurs as it grows in length, and the intronic sequences are removed. Finally, termination occurs where transcription by RNAPII and the elongation complex continues past the *polyadenylation signal*, where polyadenylation factors assemble and catalyze cleavage of the transcript. The processed transcript is now prepared for export from the nucleus to the ribosome for translation.

As described, transcription and processing of mRNA, although two different processes, have been observed to occur concurrently and it has been substantiated that they are in fact dependent on one another, allowing the nascent transcript to be processed as transcription completes and increasing the efficiency of gene transcription.

3.3.2 Translation

Translation is the process by which a protein is synthesized by the assembly of amino acids into polypeptide chains according to the “code” of the mRNA sequence. This synthesis requires interaction between mRNA, transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and other proteins that form complexes. Translation occurs on structures in the cell called ribosomes that are composed of rRNA molecules and proteins. The mRNA is attached to the ribosome via the

5′ cap and is read in a 5′-to-3′ direction. The amino acids to be added are coded by codons of three base pairs in the mRNA sequence. These mRNA codons are not directly recognized by amino acids; in fact, tRNAs molecule act as adaptors, recognizing both mRNAs and amino acids and bringing them together. In the ribosome, tRNAs align an amino acid, which is attached to the 3′ end of the molecule, corresponding to the codons of the mRNA by virtue of complementary base pairing with the anticodon loop of the tRNA.

This codon-anticodon base pairing and the attachment of the correct amino acid at the 3′ end ensure accurate amino acid incorporation in the resulting protein molecule. The ribosome scans along the mRNA to which it is bound, and translation begins at specific sites, also known as initiation sites, coding for the amino acid methionine (most commonly the AUG codon). In the ribosome structure, there are three sites available for tRNAs to bind: the peptidyl site (P site), the aminoacyl (A site), and the exit site (E site). The first tRNA, the initiator, binds to the P site; the next tRNA is then brought to the A site to bind by an elongation factor that is attached to a guanosine triphosphate (GTP) molecule.

As the tRNA is incorporated into the A site, the GTP on the elongation factor is hydrolyzed to guanosine diphosphate (GDP) and released. Subsequently, a peptide bond forms between the initiator tRNA and the second tRNA. The methionine is thus transferred to the second aminoacyl tRNA. The ribosome then reads three bases along the mRNA molecule, moving the second tRNA to the P site and the uncharged tRNA, which no longer carries an amino acid, to the E site. The A site is now available for the next tRNA to pair via codon-anticodon interaction and for the reaction to occur again, thus incorporating the next amino acid. Elongation stops when a stop codon in the mRNA is reached (UAA, UAG, or UGA) because there are no tRNAs with anticodons that pair with such codons. Instead of tRNA binding, a release factor recognizes the site. As a consequence, the mRNA and tRNAs are released from the ribosome, the completed chain of amino acids also dissociates, and the translation sequence can begin once more. (See [3], *Essential Medical Genetics*.)

3.4 GENE REGULATION

All nucleated cells of an individual have an identical genome, yet at any one time in many cell types only a small fraction of the total genes are expressed at a significant level. Moreover, during the differentiation of cells and tissues, the relative pattern of expression needs to vary widely and permit the synthesis of different proteins in each cell. These demands may change, for instance, in development and in the cellular response to local growth factors or circulating hormones. Areas of DNA flanking each gene play an important role in regulating transcription and hence synthesis of each protein. Immediately upstream of the gene is the promoter, which is involved in the binding of RNA polymerase II to the DNA template strand. Promoters for RNA polymerase II are usually several hundred nucleotides long and often contain a consensus sequence, 5′-TATAAA-3′ (the TATA box). This sequence, usually located around 25 bp upstream of the transcriptional start site, binds, via the so-called TATA-binding protein (TBP) subunit of TFIID, to a series of general transcription factors (e.g., TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH), which are relatively abundant proteins used to initiate the transcription of nearly all mRNAs.

Basal transcriptional efficiency is commonly aided by the presence of other upstream short-consensus regulatory sequences in the promoter region. These often include multiple copies of the so-called GC box, which binds the ubiquitous transcription factor SpI and the CAAT box, which is typically located 75 bp upstream of the transcriptional start site and to which the transcription factors CTF and CBF bind.

The activity of many promoters is modulated (increased or decreased) by one or more enhancers—generally short (less than 20–30 bp) sequences that bind *specific (i.e., tissue-restricted) transcription factors*. Most enhancers can function whether on the coding or the noncoding strand of DNA and can be located, in either orientation, up to several hundred kilobases from their target promoter. Most of them are active only in specific cell types and thus play a central role in regulating the tissue specificity of gene expression. Silencers are similar sequence elements that, in contrast, inhibit transcription of the associated gene.

There are several chromosomal deletions that can deleteriously affect developmental gene expression and thus cause clinical syndromes despite being located far upstream or downstream of the important genes. The presence of intergenic enhancers and silencers is likely to explain, in many cases, the pathogenicity of such deletions. Large numbers of specific transcription factors and their interaction at individual enhancers allow complex patterns of gene activation, such as during tissue differentiation and with physiological (or pathological) receptor signaling. Interactions between the proteins bound to enhancers and those bound to the promoter sequences are probably permitted by the intervening DNA's formation of long loops.

Approximately 80% of genes are expressed only at specific times and places (e.g., insulin in the pancreatic islet β-cells). The remaining 20%, called housekeeping genes, are expressed in all tissues, generally fulfill basic metabolic needs, and account for 90% or more of the genes expressed in any particular cell type. The majority of housekeeping genes and about 40% of tissue-specific genes possess CpG islands near their 5′ ends. CpG islands are about 1 kb in length and contain a high proportion of 5′-CG-3′ dinucleotide pairs, generally indicated as "CpG" (i.e., "C" for cytosine, "p" for phosphate, and "G" for guanine) to distinguish

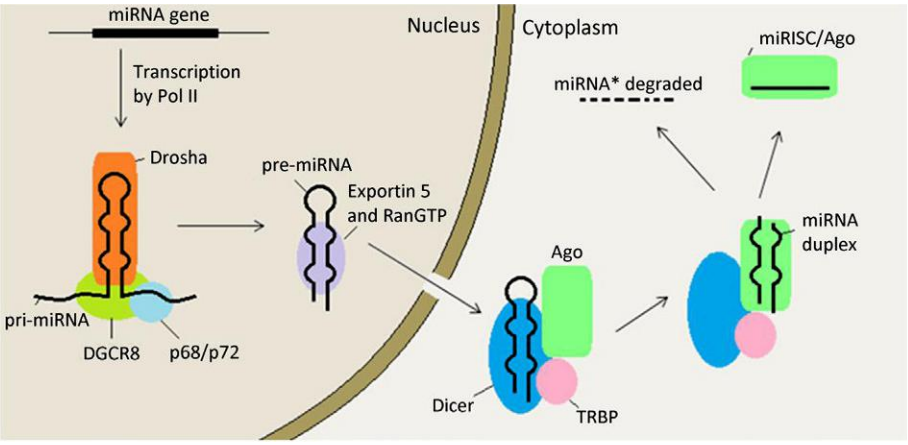


FIGURE 3.3 **Schematic of miRNA biogenesis.** Pri-miRNA transcripts are subject to two cropping events. First, Drosha, in the microprocessor complex, cleaves the mRNA-like pri-miRNA to give a pre-miRNA structure. This structure is recognized by exportin-5 and, together with RanGTP, transports the pre-miRNA to the cytoplasm via nuclear pores. In the cytoplasm the second cleavage event occurs and is catalyzed by Dicer in the RLC. This generates a miRNA: miRNA* duplex that is separated by Ago protein. The miRNA* strand or passenger strand is degraded; however, the guide strand is loaded into Ago, and other members of the RISC complex are recruited. miRISC carries out miRNA-mediated silencing of target mRNAs. *Source: Adapted from* [7] *with information from Kim et al.* (2009).

them from G hydrogen-bonded to C in opposite DNA strands. Generally, cytosine residues in these CpG dinucleotides are methylated. There is usually a lack of methylation, however, in CpG islands associated with active neighboring gene expression.

Although the bulk of gene expression regulation probably occurs at the transcriptional level, the regulation of a gene's activity may also occur later. For example, there may be (1) alternative splicing of the mRNA to produce different gene products (e.g., calcitonin or calcitonin gene–related neuropeptide) or (2) an altered rate of translation or degradation of the protein.

Mutations in a gene's regulatory sequences (such as its promoter region) can occur and may result in no gene product (e.g., some patients with β-thalassemia), abnormal continued expression of a fetal gene product (e.g., hereditary persistence of α-fetoprotein or hemoglobin F), or anomalous patterns of gene expression (e.g., ectopic expression of creatine kinase).

There are now important internationally agreed conventions regarding the nomenclature, for publishing and clinical reporting, of human gene mutations and their effects on proteins (see [3]).

3.4.1 Gene Regulation by miRNAs

MicroRNAs (miRNAs) are small noncoding RNAs of ~22 nt that play an essential role in gene regulation by binding to the 3′ UTR of protein-coding mRNAs, and by initiating post-transcriptional silencing by repressing translation and/or causing mRNA degradation. There are currently 2578 mature miRNAs reported in miRBase. As each miRNA can have multiple targets, collectively they are thought to directly regulate approximately a third of human genes (see [8,9]). There are high levels of conservation in miRNAs between organisms, particularly bilaterian animals, suggesting that they play a vital role in multiple species. The expression pattern of miRNAs varies between cell types and over time; thus, it implicates them heavily in development and other biological processes. Additionally, deregulation of miRNA expression has been shown to have pathological consequences.

Genes encoding miRNAs are located throughout the genome, although many are found situated close together under the control of a single promoter region. Normally, miRNAs or miRNA clusters are found in introns of protein-coding genes or in noncoding DNA; however, some have been found in exons of protein-coding genes that rely on alternative splicing for their expression. miRNA gene transcription is usually carried out by RNA polymerase II (Pol II), although in rare cases Pol III is used. This places the majority of miRNA genes under the control of transcription factors associated with Pol II, along with many other transcription factors, and thus Pol II controls the expression of miRNAs in different cell types and during development.

Primary transcripts (pri-miRNA) must undergo processing before the mature functional miRNA can bind target mRNAs (Figure 3.3). The pri-miRNA structure has a 5′ cap and a poly(A) tail, and it usually contains at least one long (~30 bp stem) imperfectly matched hairpin loop with single-stranded segments on either side of its base (see [7,10]). The pri-miRNA structure is recognized by the microprocessor complex, which consists of an RNase III enzyme referred to as Drosha, and an RNA-binding protein referred to as DiGeorge syndrome critical region gene 8

(*DGCR8*). Additional co-factors such as RNA helicases p68 and p72 are required for biogenesis of specific miRNAs. DGCR8 recognizes the junction between the single-stranded RNA and the double-stranded hairpin and acts to guide Drosha to cleave ~11 bp away from it. The truncated hairpin product of this reaction is known as precursor miRNA (pre-miRNA) and is characterized by a 3′ overhang created by Drosha cleavage. Microprocessor activity is thought to occur co-transcriptionally and before splicing of the coding RNA, but does not inhibit splicing activity.

Exportin-5 recognizes the 3′ overhang and the double-stranded stem of the hairpin and binds pre-miRNAs. Together with Ran-GTP, it exports pre-miRNA from the nucleus to the cytoplasm. Once in the cytoplasm, pre-miRNA dissociates from exportin-5 and is recognized by the RISC loading complex (RLC), which consists of a second RNase III enzyme, Dicer, TAR RNA-binding protein (TRBP), and members of the argonaute (Ago) protein family. Dicer recognizes the 3′ end of the pre-miRNA and cuts both strands of the hairpin ~22 nt away from the 5′ and 3′ ends, thus removing the terminal loop and creating an miRNA duplex. TRBP function is not yet fully understood, although it is possible that it stabilizes the RLC or plays a role in proper RNA-induced silencing complex (miRISC) formation. Only one strand of miRNA is loaded onto an Ago protein to form part of the RISC, while the other strand, termed miRNA*, is degraded. The RISC complex facilitates the miRNA-mediated repression of target mRNAs, which will be discussed in more detail later.

The properties that determine which strand enters the miRISC are not fully elucidated. Thermodynamic stability of the ends of both miRNA strands and the positions of mismatches in the duplex are thought to play a role; however, there are cases where both strands are used as mature miRNAs and where the relative frequency of the strand utilized is similar. Heterogeneity also affects the frequency of specific miRNAs. It is brought about by Drosha and Dicer cleavage activity that is not absolutely specific, thus generating strands of differing lengths and ends. The specificity of the miRNA for its target can be altered if there is a change at the 5′ end of the miRNA because this can affect the seed. The seed is the region of the miRNA that binds specifically to targets in the 3′ UTR of mRNA and is typically situated in nucleotides 2–7 of the miRNA.

The processes just described form the canonical miRNA biogenesis pathway; however, there are noncanonical methods of miRNA maturation. Mirtrons are spliced from the transcribed intron of an mRNA and directly form a pre-miRNA structure, thus making the microprocessor stage of biogenesis redundant, but the relevance of this type of miRNA is not yet understood. There also exists at least one miRNA, miR-451, that can skip the Dicer step in maturation. Pre-miR-451 can directly bind Ago2 and is cleaved in a two-step process to form miR-451 (see [7]).

3.4.2 mRNA Targets

In plants the full length of a miRNA binds to its target in mRNA molecules, with pairing at every, or nearly every, base. However, in animals full complementarity is an infrequent occurrence. Instead, as mentioned previously, the seed (nucleotides 2–7) of the miRNA is the region found in complete or almost complete Watson-Crick pairing with the mRNA's target site. Targets are located in the 3′ UTR in animal mRNA and are nonrandomly distributed; the majority of are located at the 3′ and 5′ ends of the 3′ UTR. In addition, multiple target sites are often found in one 3′ UTR, and it is thought that miRNA binding at more than one site can act to further increase mRNA repression (see [11]).

While seed pairing is essential, the level of complementarity can vary slightly depending on the target sequence; thus, the different interactions can be classified as canonical, 3′-complementary, and 3′-supplementary (Figure 3.4). Canonical sites are naturally the most abundant and conserved site types. First is 7mer-1A, which has seed pairing at positions 2–7 with an adenine opposite position 1 of the miRNA. The adenine increases the repression by miRNAs. Second is 8mer, which has the same composition as 7mer-1A but with an additional pairing at position 8 of the miRNA. Third is 7mer-m8, which has pairing with the seed at positions 2–8 but does not have an adenine opposite position 1. Both the 3′-supplementary and 3′-compensatory sites involve pairing at downstream 3′ sites in the miRNA (nucleotides 13–16). Where the 3′ site is additional to normal 7mer/8mer binding, it is supplementary; however, this additional pairing has been shown to have little effect on the efficiency of silencing. The rarely seen compensatory 3′ pairing is so-called because it compensates for a 1-base mismatch in the seed region. Although their silencing effect is minimal, 6mer sites can also been found.

Bartel [12] assessed different factors in interactions between miRNA and mRNA that affect repressive action. This study found that 8mer sites are the most effective of all basic sites, while multiple sites in a single 3′ UTR have the greatest overall efficacy. It was also reported that increased AU content of the region surrounding the target site is directly correlated with repressive effect.

Finding target mRNAs is important for understanding the cellular processes in which specific miRNAs are involved. Generally, computational methods are used to identify possible candidate targets that are then verified experimentally. Computational target prediction is complicated by the small and variable region of complementarity between miRNAs and their targets in animals (see [13]). Almost all such programs rely on seed complementarity in some form for selection of targets from 3′ UTR datasets. However, given the small size of the seed region, this complementarity produces an overwhelming number of results and so a number of programs incorporate additional

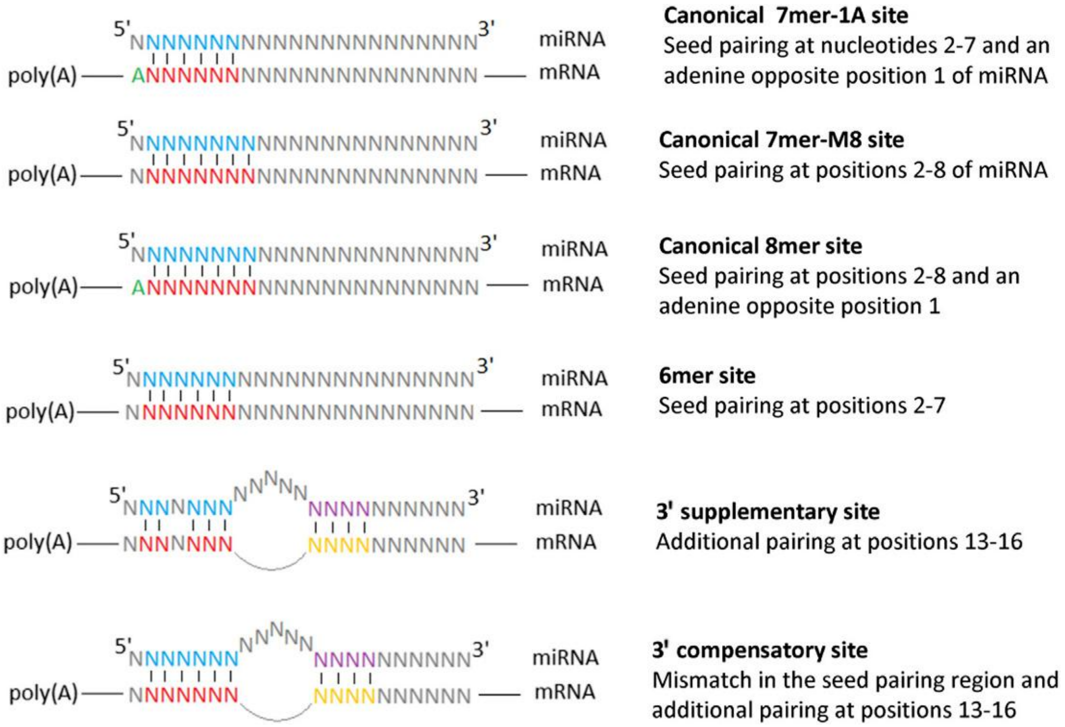

5' NNNNNNNNNNNNNNNNNNNNN 3' miRNA
poly(A) — ANNNNNNNNNNNNNNNNNNNN — mRNA
Canonical 7mer-1A site
Seed pairing at nucleotides 2-7 and an adenine opposite position 1 of miRNA
Canonical 7mer-M8 site
Seed pairing at positions 2-8 of miRNA
Canonical 8mer site
Seed pairing at positions 2-8 and an adenine opposite position 1
6mer site
Seed pairing at positions 2-7
3' supplementary site
Additional pairing at positions 13-16
3' compensatory site
Mismatch in the seed pairing region and additional pairing at positions 13-16

FIGURE 3.4 **Classification of miRNA target sites.** The blue bases represent the seed region and the red bases are the complementary sites in the target mRNA. The adenine found opposite miRNA position 1 is depicted in green. Binding 3′ of the seed region is shown in purple and orange bases and occurs at bases 13–16 of the miRNA, forming a bulge between the seed-paired regions and the 3′-paired region. *Source: Redrawn with modifications from* [11].

parameters to narrow the search. These criteria include conservation of the site between organisms, the presence of more than one target in close proximity, the structure of the 3′ UTR, and the thermodynamic stability of the miRNA–mRNA interaction. As each of these programs has its own limitations regarding specificity and sensitivity for target detection, it is often suggested that combinations of more than one algorithm be used to produce more reliable results. The expression patterns of miRNAs and the selected targets should then be compared using array technology; targets that show an inverse correlation with the level of miRNA should be experimentally validated (see [11]).

3.4.2.1 *Mechanisms of Silencing*

Where there is almost full complementarity between miRNAs and mRNAs, Ago2-mediated cleavage of the target mRNA can occur. However, as discussed, full complementarity is rare, and so translational repression was originally thought to be the silencing mechanism employed by animal miRNAs (see [14]). Nevertheless, studies on mRNA levels have shown that, in coordination with miRNA depletion, there is an increase in target mRNAs. Likewise, an increase in miRNA levels results in a decrease in target mRNAs [14]. It has now been shown that mRNA degradation may in fact be the predominant mechanism by which silencing is carried out. Currently, there is still much debate as to whether translational repression or mRNA destabilization is the initial event triggered by miRNA-facilitated miRISC activity. Furthermore, the exact alterations and reactions that are catalyzed by the miRISC and that cause silencing are yet to be determined. Despite this, some key components have been identified and will be briefly discussed.

As mentioned previously, after the mature miRNA is loaded into an Ago family protein, it forms part of the miRISC. Another member of this complex is TNRC6, which contains GW repeats along the length of the protein and has been shown to associate with Ago through its N-terminal region (see [15]). TNRC6 has multiple paralogs (TNRC6A, -B, and -C) and is conserved across species—for example, GW182 in *D. melanogaster*. Studies have assessed the importance of various domains in TNRC6 with regard to mRNA silencing and its interactions with other proteins (Figure 3.5). The C-terminal silencing domain (SD) is essential for silencing and contains sites for interaction with cytoplasmic poly(A)-binding protein (PABP); NOT1, which is a member of the CCR4-NOT complex; and PAN3,which is part of the PAN2-PAN3 complex [15].

The CCR4-NOT complex is a deadenylase, shown to bind TNRC6 through NOT1 interaction with tryptophan repeats found in M1, M2, and the C-terminal. PAN2-PAN3 is also a deadenylase and interacts with the M2 domain in

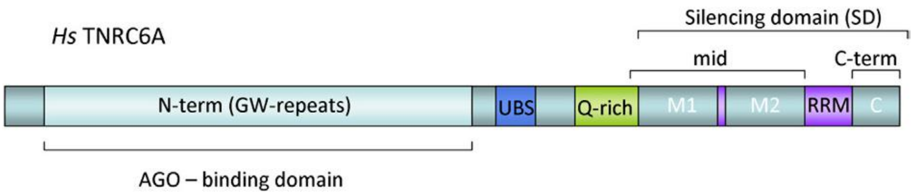


FIGURE 3.5 **Functional domains of the human TNRC6 protein.** UBA: ubiquitin associated domain. Q-rich: glutamine-rich region. M1/M2: two sections of the mid-region separated by PAM2. PAM2: PABP interacting motif 2. RRM: RNA recognition motif. *Source: Drawn from* [14].

TNRC6. Deadenylation of the poly(A) tail of target mRNAs leads to deadenylation-dependent mRNA degradation by exonucleic action. The role of CCR4-NOT in mRNA degradation has been confirmed by experiments showing that the tethering of this complex to target mRNAs, thus bypassing the need for miRNA-mediated interactions, nevertheless causes deadenylation and subsequent mRNA degradation (see [15,16]). The CCR4-NOT complex has also tentatively been implicated in translational repression, as tethering to mRNA targets without a poly(A) tail still allows for silencing without deadenylation-driven degradation. Conversely, removing subunits of CCR4-NOT leads to some decrease in the silencing effect on mRNAs without a poly(A) tail. Other studies have uncovered evidence against CCR4-NOT participation in translation repression, showing that its depletion still allows for some silencing by translation repression but that mRNA degradation is inhibited (see [15]).

PABP binds to mRNA poly(A) tails and also associates with eukaryotic translation-initiation factor 4F (eIF4F) complex and the 5′ cap to generate a circular mRNA structure that is stable and efficiently translated. TNRC6 can also interact with PABP via the PAM2 domain, and disruption of this interaction leads to a decrease in the silencing ability of miRISC, even when TNRC6 continues to interact with deadenylases [15]. The interaction with TNRC6 causes PABP to disassociate from the poly(A) tail before the initiation of deadenylation and at the same time that Ago association with mRNA is observed (Figure 3.6; see [17]). This suggests that miRISC coordinates the removal of PABP from the poly(A) tail and allows for CCR4-NOT deadenylation. PABP has also been implicated in the recruitment of miRISC to target miRNAs. The fact that the presence of a poly(A) tail enhances silencing but is not strictly necessary for it indicates that, while PABP plays an important role, it may not be essential for silencing under all conditions.

The removal of PABP from the poly(A) tail may also disrupt EIF4F complex binding and thus negatively affect translation initiation. A study by [18] showed that the EIF4F complex is required for repression of mRNAs with a structured 5′ cap. Additionally the study assessed the specific component involved in EIF4A2 by knockdown of each member of the complex. EIF4A2 is thought to unwind the 5′ secondary structure of mRNA for translation. Meijer et al. [18] showed that EIF4A2 is bound to repressed mRNAs and also immunoprecipitates with NOT7, a member of the CCR4-NOT complex. EIF4A1 is normally found in excess of the A2 form and is thought to be involved in normal translation activity, whereas repression-induced binding of EIF4A2 inhibits translation. However, this interaction should be further investigated to confirm the role of this protein in translational repression.

While it is obvious that the precise mechanisms of both mRNA degradation and translational repression are as yet unclear, key components in both pathways have been identified. Huntzinger et al. [15] proposed that the minimal requirements for repression are Ago, TNRC6, PABP, and the deadenylation complex, and that other factors may regulate or enhance the repression activity.

3.4.3 Chromatin Modifications and Gene Expression

Particular residues present in histone tails are often subject to post-translational modification. For example, the basic residues lysine (with mono-, di- or tri-methylation), arginine (mono-, symmetrically, and asymmetrically di-methylated), and histidine are commonly methylated (see [6,19]), whereas serine, threonine, and tyrosine are often phosphorylated. Depending on the position of the modification, this can have different effects on overall chromatin structure; for example, the tri-methylation of lysine 27 on histone 3 (H3K27me3) is associated with the closing of the chromatin, and therefore transcriptionally silent, regions. However, despite methylation being associated with silencing, this is not always the case. Methylation of lysine 4 on H3 (H3K4) has been observed to have the opposite effect. The many types of modifications and the histone tail residues with which they are associated are summarized in Figure 3.7.

3.4.4 Nucleosome Positioning and Occupancy

For the binding of TFs and other factors essential for transcription to take place, regions of heterochromatin must be remodeled and the nucleosome positions moved [20]. The positioning of nucleosomes controls many DNA

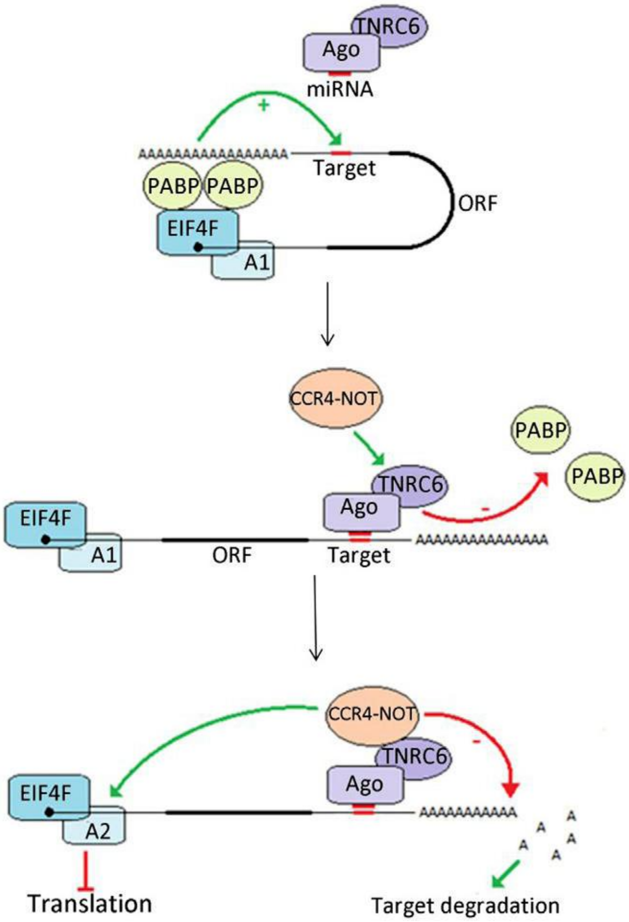


FIGURE 3.6 **Proposed mechanism of mRISC-mediated mRNA destabilization and translational repression.** PABP is bound to the poly(A) tail of the target mRNA and is associated with the 5′ cap through the EIF4F complex, creating a loop that is stable for translation. PABP aids in recruiting mRISC via TNRC6 binding. It dissociates from the poly(A) tail upon mRISC binding to the target, causing the mRNA loop to break. TNC6 recruits the CCR4-NOT complex by binding NOT1. CCR4-NOT causes deadenylation of the poly(A) tail, which destabilizes the mRNA and triggers exonuclease-mediated mRNA degradation. Additionally, mRISC binding causes EIF4A2 to bind the 5′ cap instead of the A1 isoform, which inhibits translation of the mRNA.

interactions, including determining which sequences can undergo processes such as transcription or be repaired (see [21]). Nucleosomes are not distributed evenly across the genome; some sequences are more likely to have dense nucleosome occupancy, while others (such as enhancer, promoter, or other regulatory sequences rich in poly A: T and C: G tracts) tend to have fewer such structures.

Some sequences have even been observed to be completely nucleosome-free. Other factors, such as DNA sequence, can influence the position of nucleosomes in particular regions (Figure 3.8). For example, more flexible DNA sequences rich in A and T dinucleotide interspersions (about every 10 bp) tend to come into contact with the histone octamer because they can bend around the protein core while more rigid sequences do not come into as close contact with the complex. The linker DNA regions interspersed between nucleosomes are also associated with more rigid sequences such as poly(A) tracts. Both positioning and occupancy determine which DNA interactions can take place. It is of note that nucleosomes inhibit protein interactions with DNA, although to varying extents depending on the interacting factor itself. For example, some TFs may bind to DNA even in the nucleosome structure, whereas other factors are incapable of initially interacting with DNA where nucleosomes are present without the aid of remodeling complexes.

There are two main mechanisms by which chromatin structure can be altered: the first, through covalent histone modifications by specific enzymes; the second, through ATP-dependent remodeling complexes. In the cellular context, these mechanisms work together in chromatin structure regulation.

3.5 GENOMIC VARIATION

Genetic differences in individuals within and between populations are responsible for human diversity, including traits such as eye and skin color and even disease susceptibility.

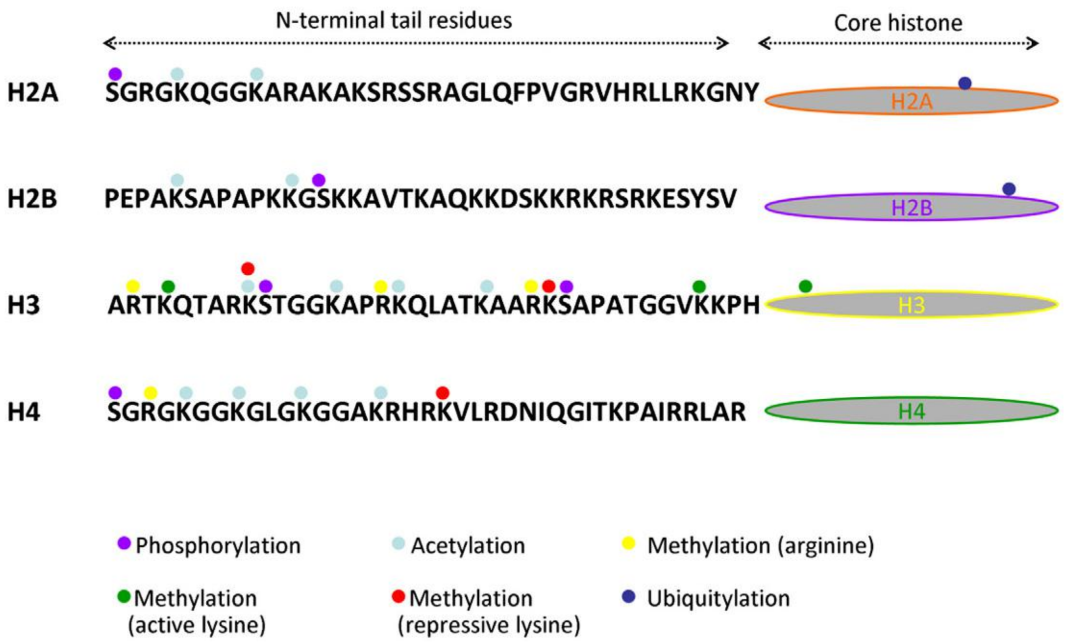

N-terminal tail residues
Core histone
H2A
SGRGKQGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGNY
H2A
H2B
PEPAKSAPAPKKGSKKAVTKAQKKDSKKRKRSRKESYSV
H2B
H3
ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPH
H3
H4
SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLAR
H4
Phosphorylation
Acetylation
Methylation (arginine)
Methylation (active lysine)
Methylation (repressive lysine)
Ubiquitylation

FIGURE 3.7 **Post-translational chemical modifications to histone proteins.** Phosphorylation of serine residues (purple dot) and acetylation of lysine residues (blue dot) occur on all four histones of the nucleosome. Ubiquitylation occurs on histones H2A and H2B. Methylation of arginine residues occurs on H3 and H4. On H3, many lysines are methylated. However, whether the methylation causes repression or activation depends on which lysine is modified; for example, H3K4 and H3K36, when methylated, are associated with activation, whereas methylation of H3K9 and H3K27 is implicated in silencing. *Source: Drawn using information from Allis et al.* (2006) [22].

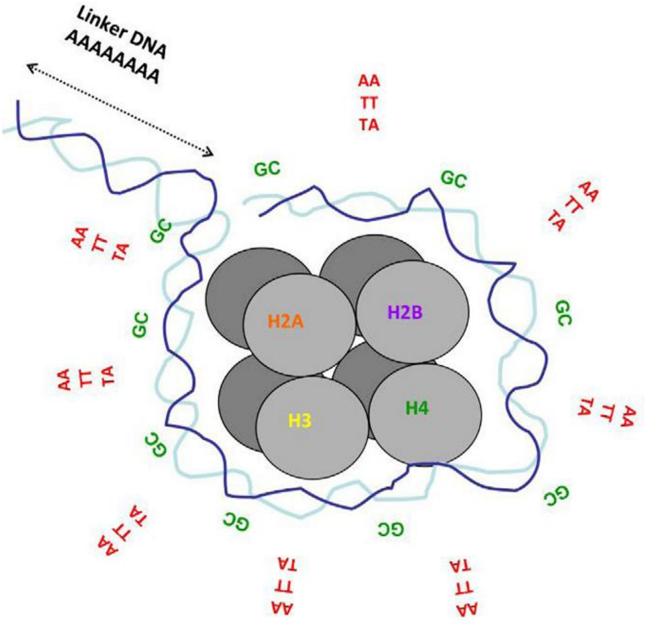

Linker DNA
AAAAAAAA
AA TT TA
GC
H2A
H2B
H3
H4

FIGURE 3.8 **DNA sequence preferences of the nucleosome.** *Source: Redrawn with modifications from* [21].

Intra- and interpopulation variation has arisen over time as a result of, for example, mutation, recombination, and natural selection. Natural selection may occur, for instance, when alleles or sequences confer selective advantages for individuals carrying such variants—in response to, say, a particular environmental stress to which their population has been exposed. These allele frequencies vary within populations and between them. Different types of variation are found in the genome, ranging from cytogenetically visible changes involving whole chromosomes (e.g., trisomy 21) to single base pair alterations (single-nucleotide polymorphisms). Variations in genome structure and sequence can cause phenotypic variation, both pathogenic and nonpathogenic, by altering expression levels, protein function, or protein efficacy. Genetic diversity in terms of recently discovered submicroscopic variations will be discussed in this section.

3.5.1 Single-Nucleotide Polymorphisms

Single-nucleotide polymorphisms (SNPs) are single base pair substitutions that mostly result from replication errors not repaired by DNA repair machinery, although they can occur by other means. SNPs can be located throughout the genome, but they are more frequently found in noncoding regions than in coding regions because of selective pressure on protein-coding domains. This means that alleles that confer an advantage are usually conserved.

3.5.1.1 Effect of SNPs on Gene Expression

SNPs in a coding region can be synonymous; that is, that they do not result in a change in amino acid. Alternatively, they can be nonsynonymous SNPs of two types. Missense mutations cause a substitution of a single amino acid, whereas nonsense mutations change an amino acid to a stop codon, thus truncating the protein. Type and location can determine the level of phenotypic effect that a SNP produces, with nonsense mutations more likely to be pathogenic and synonymous SNPs more often neutral [23]. The effect of missense mutations depends on the location of the altered amino acid in the protein as well as the nature of the substitution.

SNPs in noncoding regions may still affect gene expression, particularly if located in a regulatory region, as this can affect the binding of regulatory subunits or transcriptional machinery (see [23]). Also, SNPs located in introns may cause variants by altering gene splicing. Many small non-coding regions of the genome are transcribed and affect post-transcriptional regulation of gene expression—namely, small noncoding RNAs (as described in the section on gene regulation). SNPs in these regions or their target regions in mRNAs can also cause variation in gene expression.

3.5.1.2 Linkage and Haplotype Blocks

Variation in the SNP profile of individuals and populations is brought about by recombination, which is the exchange of material between chromosomes that can occur during meiosis when crossovers are formed between homologous regions (Figure 3.9). Linkage describes the principle that loci closer together in the genome are more likely to be inherited together because there is less chance of a recombination event occurring between them. A haplotype is a set of alleles inherited together. Generally, recombination will lead to new haplotypes, and at similar frequencies between equally spaced loci. However, linkage disequilibrium leads to some alleles remaining together through multiple generations because recombination does not occur at the expected frequency. This can be due to selective advantage or to one

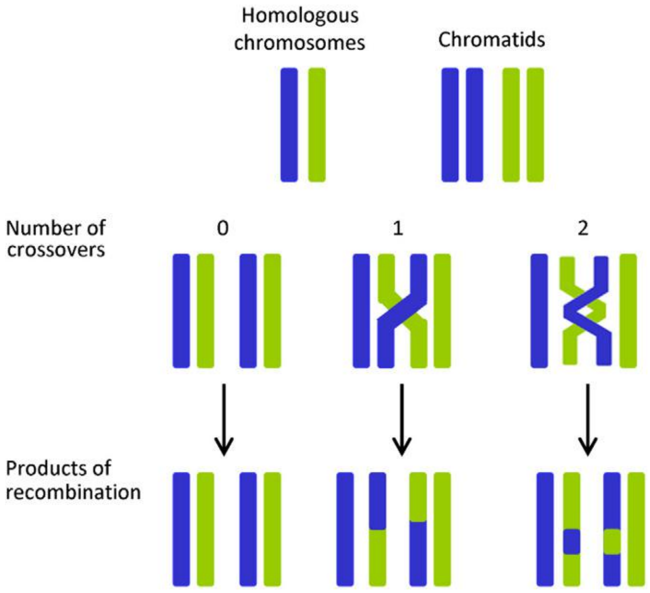


FIGURE 3.9 Recombination during meiosis. Pairs of chromosomes are held together during prophase I of meiosis by crossovers that form between homologous sites. When these pairs are pulled apart, there can be an exchange of material, causing a new combination of alleles to be generated on one chromatid of each pair. After meiosis II, gametes with four haplotypes are formed.

allele being relatively new and so there has been insufficient time for recombination to occur (see [3]).

There are specific locations in the genome that are more prone than surrounding regions to recombination events. As a result, blocks of sequence between these hotspots tend to exhibit linkage disequilibrium and are inherited as intact units, termed haplotype blocks. For each haplotype block, there can be multiple haplotypes, each containing a different SNP combination.

The size and frequency of haplotype blocks vary between populations, and this can be used in population origin studies. The greatest amount of haplotype diversity is seen in African populations. In contrast, repeated founder effects during migrations have led to less haplotype variation in other populations. Larger conserved blocks indicate a more recent ancestry because there has been less time for recombination to occur; again, this is consistent with the generally shorter nature of blocks seen in African populations, and it may also be due to selective advantage conferred on a haplotype. Haplotype blocks can be useful in disease studies, as mutations located in a SNP haplotype block can be associated with a particular haplotype. Thus, SNPs can be used to narrow the search for the block that contains the disease-causing mutation, by assessing the frequency with which key identifying SNPs (tag-SNPs), in different blocks associate with the disease.

3.5.2 Copy Number Variants

Copy number variants (CNVs) are not as well characterized as SNPs, but they are believed to be, in many cases, functionally significant in human disease and variation. CNVs are a form of structural variation that arises from dosage effects (relative numbers of each sequence) or potentially from the interruption of sequences. They are distinct from SNPs as they range in size from kilobases to megabases. Their occurrence is due to events such as deletions, duplications, or insertions in particular regions known to be spread throughout the whole genome. In fact, in a study of 270 individuals of European, Asian, and African ancestry, 1447 CNV regions were found by Redon et al. (2006) [24], including associations with genes, disease loci, and functional elements that contribute to transcription. Particular CNVs are also known to be associated with human diseases such as Parkinson's and Alzheimer's.

3.5.2.1 Mechanisms of CNV Formation

The sequence context of CNVs has been examined for clues about the mechanisms involved in the formation of CNV regions. Homologous recombination is believed to be one mechanism, occurring between sequences that are very similar but not allelic. Duplications of near identical (>90%) sequences, known as segmental duplications, as well as deletions, have been observed and are thought to have arisen in this way. Other genomic sequence homologies, such as Alu elements (sequences dispersed across the whole genome) are also believed to play a role in the promotion of nonallelic homologous recombination. Orientation, size, and sequence similarity are factors expected to have an effect on the likelihood of such mechanisms occurring and causing rearrangements. The outcomes of combinations of these mechanisms and factors can be seen in Figure 3.10.

3.5.2.2 Population Genetics of CNVs

Data on particular CNVs found in the population studies carried out by Redon et al. (2006) show that variants are often population-specific. The distribution of CNVs within and between populations is determined by mutation and selection among other factors.

The human salivary amylase gene *(AMY1)* is one of the most polymorphic of the human CNVs, with the number of copies of it varying from 6 to 15. Such an adaptation is believed to have arisen and subsequently been selected for because of high-starch diets in some populations, improving the breakdown and digestion of starch by amylase into maltose sugar.

3.5.2.3 Relationship of CNVs and Transcriptional Regulation

The majority of CNVs, like SNPs, are found in noncoding regions of the genome or in relatively gene-sparse regions. Changes in copy number can have an effect on gene activity directly, in a relative dosage–sensitive manner, or indirectly, by interrupting sequences that are essential in transcriptional regulation—for example, acting as activators or repressors of transcription (Figure 3.11). Thus, as with other forms of genetic variation (e.g., SNPs), differences in protein levels that arise from CNVs by direct or indirect mechanisms may disrupt cellular or tissue homeostasis and may have a role in human disease. Conversely, the threshold level of change that is required to confer a phenotypic effect may not be reached by these mechanisms, and thus some CNVs may result in no observable effects in an individual or on those carrying the variant [25].

The study of CNVs and their effects on human variation and disease is still very much in its infancy; however, from the findings so far it is likely that their impact on our understanding of variation and, potentially, the molecular mechanisms that confer risk to disease phenotypes will be highly significant [26].

3.5.3 The ENCODE Project and the Characterization of Human Variants

The Encyclopedia of DNA Elements (ENCODE) project, initially launched in September 2003, is a research group

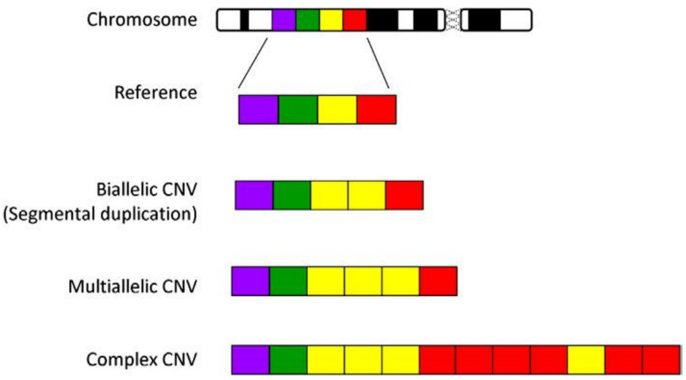

Chromosome
Reference
Biallelic CNV
(Segmental duplication)
Multiallelic CNV
Complex CNV

FIGURE 3.10 **Copy number variation.** Purple and green segments are constant. Yellow and red segments differ in number in relation to the other segments, resulting in different CNVs: segmental duplication, multiallelic variants, and more complex variants.

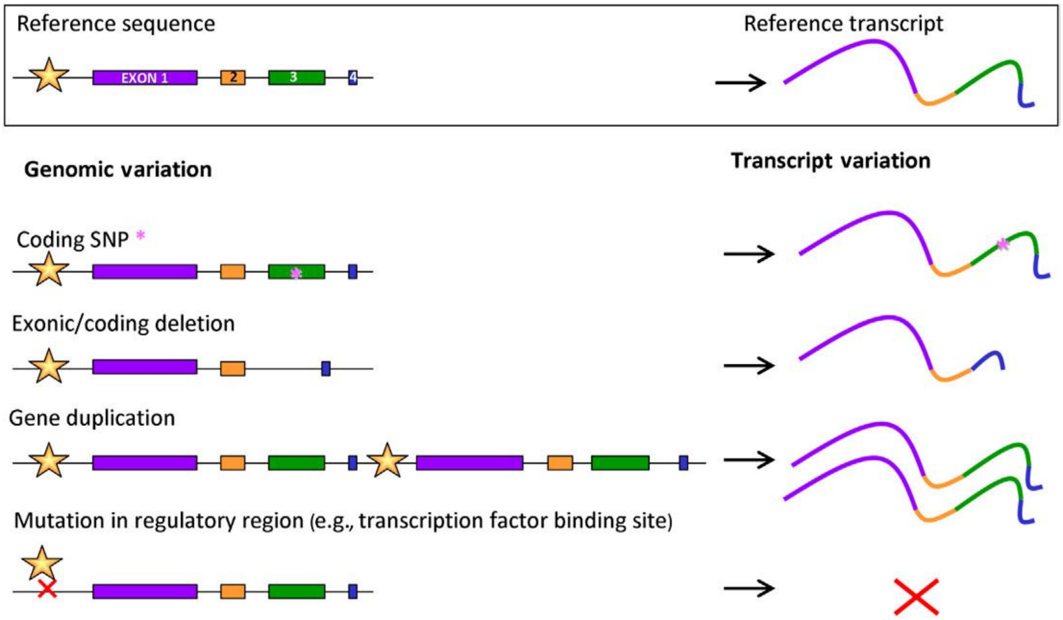

Reference sequence
EXON 1
2
3
4
Reference transcript
Genomic variation
Transcript variation
Coding SNP *
Exonic/coding deletion
Gene duplication
Mutation in regulatory region (e.g., transcription factor binding site)

FIGURE 3.11 **Relationship between genome variation types and transcripts produced.** Coding variations such as SNPs, deletions, or duplications are responsible for transcriptional alterations at the sequence level. Variations via regulatory mechanisms are responsible for changes in transcript abundance. *Source: Redrawn with modifications from* [23].

working under the auspices of the National Human Genome Research Institute (NHGRI). Its goal is to identify regions of the genome that are active or that have functional significance. So far, ENCODE has estimated that, as much as 80.4% of the genome may have biochemical (and possibly also biological) activity, using evidence from at least one cell type (out of the 147 cell types used) and using at least one experimental or computational system. Specifically, it is reported that 8.5% of the genome consists of protein-binding sites for transcription factors. ENCODE researchers have also calculated that the total number of enhancer elements in the human genome, according to their data, is 399,124 and that there are, on average, 6.3 alternative transcripts per locus.

Many investigations of complex or multifactorial disorders have involved genome-wide association studies identifying a large number of SNPs conferring disease susceptibility that lie outside the coding regions of genes. Variants that are associated with human diversity (SNPs, CNVs, epigenetic changes) have also been characterized.

The enormous volume of work undertaken for the ENCODE project has resulted in many published papers and datasets. These data and published conclusions (particularly concerning the proportion of the genome calculated to be biologically functional; see [27]) have provoked much interesting debate and even some considerable controversy internationally.

3.6 CONCLUSION

There are multiple complex and dynamic mechanisms, including histone modification and nucleosome repositioning, by which the genome's compacted structure is maintained and modified. These mechanisms, together with the actions of RNA genes (e.g., miRNAs) and the consequences of interindividual genomic sequence variation (e.g., SNPs and CNVs) affect gene expression in ways that are incompletely understood but include altered transcription, splicing, and translation. Major international research projects, such as ENCODE, currently aim to elucidate the possible biological functions of the many noncoding regions of the genome. There are many completed and ongoing genome-wide association studies, together with an increasing abundance of high-resolution array comparative genomic hybridization and whole genome sequencing data. These datasets encompass numerous SNPs and CNVs located outside protein-coding sequences in genomic regions of uncertain biological significance [2,25,28–31]. Studies into the functions of such regions is of considerable importance, from both biological and clinical perspectives.

REFERENCES

[1] <www.essentialmedgen.com> (ESSENTIALMEDGEN: portal and guide to 70 useful worldwide genetic databases and online tools, maintained by Professor Tobias).

[2] <www.ensembl.org/index.html> (ENSEMBL Genome Browser, based at the Sanger Institute, Cambridge, England.

[3] Tobias ES, Connor M, Ferguson-Smith M. Essential Medical Genetics. 6th ed. Wiley-Blackwell; 2011.

[4] Sajan SA, Hawkins RD. Methods for identifying higher-order chromatin structure. Annu Rev Genomics Hum Genet 2012;13:59–82.

[5] North JA, Simon M, Ferdinand MB, Shoffner MA, Picking JW, Howard CJ, et al. Histone H3 phosphorylation near the nucleosome dyad alters chromatin structure. Nucleic Acids Res 2014 [advance access Febraury 21].

[6] Miller JL, Grant PA. Epigenetics development and disease. Subcellular Biochem 2012;61:289–300.

[7] Blahna MT, Hata A. Regulation of miRNA biogenesis as an integrated component of growth factor signaling. Curr Opin Cell Biol 2013;25:233–40.

[8] Koturbash I, Zemp FJ, Pogribny I, Kovalchuk O. Small molecules with big effects: the role of the microRNAome in cancer and carcinogenesis. Mutat Res 2011;722(2):94–105.

[9] Salzman DW, Weidhaas JB. SNPing cancer in the bud: microRNA and microRNA-target site polymorphisms as diagnostic and prognostic biomarkers in cancer. Pharmacol Ther 2013;137(1):55–63.

[10] Starega-Roslan J, Koscianska E, Kozlowski P, Krzyzosiak WJ. The role of the precursor structure in the biogenesis of microRNA. Cell Mol Life Sci 2011;68(17):2859–71.

[11] Witkos TM, Koscianska E, Krzyzosiak WJ, et al. Practical aspects of microRNA target prediction. Curr Mol Med 2011;11:93–109.

[12] Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009;136:215–33.

[13] Lhakhang TW, Chaudhry MA. Current approaches to micro-RNA analysis and target gene prediction. J Appl Genet 2012;53:149–58.

[14] Huntzinger E, Izaurralde E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet 2011;12(2):99–110.

[15] Huntzinger E, Kuzuoglu-Ozturk D, Braun JE, Eulalio A, Wohlbold L, Izaurralde E. The interactions of GW182 proteins with PABP and deadenylases are required for both translational repression and degradation of miRNA targets. Nucleic Acids Res 2013;41(2):978–94.

[16] Hafner M, Ascano M Jr., Tuschl T. New insights in the mechanism of microRNA-mediated target repression. Nat Struct Mol Biol 2011;18:1181–2.

[17] Moretti F, Kaiser C, Zdanowicz-Specht A, Hentze MW. PABP and the poly(A) tail augment microRNA repression by facilitated miRISC binding. Nat Struct Mol Biol 2012;19(6):603–8.

[18] Meijer HA, Kong YW, Lu WT, Wilczynska A, Spriggs RV, Robinson SW, et al. Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. Science 2013;340(6128):82–5.

[19] Greer EL, Shi Y. Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet 2012;13:343–57.

[20] Leung A, Schones DE, Natarajan R. Using epigenetic mechanisms to understand the impact of Common disease causing alleles. Curr Opin Immunol 2012;24:558–63.

[21] Struhl K, Segal E. Determinants of nucleosome positioning. Nat Struct Mol Biol 2013;20(3):267–73.

[22] Allis CD, Jenuwein T, Reinberg D, Caparros ML. Epigenetics. Cold Spring Harbor Laboratory Press; 2007.

[23] Haraksingh RR, Snyder MP. Impacts of variation in the human genome on gene regulation. J Mol Biol 2013;425(21):3970–7.

[24] Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature 2006;444(7118):444–54.

[25] <http://decipher.sanger.ac.uk/> (DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources).

[26] Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. Annu Rev Med 2010;61:437–55.

[27] ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489(7414):57–74.

[28] <http://genome-euro.ucsc.edu> (UCSC Genome Browser, based in California).

[29] <www.ncbi.nlm.nih.gov/snp> (NCBI dbSNP database of single-nucleotide polymorphisms).

[30] <http://hapmap.ncbi.nlm.nih.gov/> (HAPMAP database).

[31] <www.1000genomes.org/> (1000 Genomes Project: catalog of human genetic variation).

Chapter 4

Epigenetics

Tony McBryan and Peter D. Adams

Institute of Cancer Sciences, University of Glasgow and Beatson Institute for Cancer Research, Glasgow, UK

Chapter Outline

4.1 Overview 57
4.2 DNA Methylation 57
4.2.1 Effects of DNA Methylation 58
4.2.2 DNA Methyltransferases 60
4.2.3 Techniques for Measuring Methylation 60
4.3 Histones and Variants 61
4.3.1 Histone PTMs 63
4.3.2 Histone PTM Effects 64
4.3.3 Techniques for Measuring Histones and Other Chromatin-Interacting Proteins 64
4.4 Epigenetic Disorders And Therapies 65
4.5 Conclusion 66
References 66

4.1 OVERVIEW

In a single organism, there are a huge variety of cell types that each perform specific, and often unique, tasks. Fibroblast cells are responsible for generating the stroma in tissue. Cardiac muscle is electrically sensitive and, through coordinated contractions signaled by electrical pulse, propels blood through the blood vessel. Neuronal cells are also electrically sensitive, but electrical activity instead activates synaptic connections with other cells. Parietal cells produce a potent gastric acid in the stomach to facilitate digestion of food. All of these cells are derived from a common fertilized egg cell and all share the same DNA sequence; remarkably, however, each has a vastly different purpose and subsequently a vastly different gene expression program.

Epigenetics refers to any heritable change in genome function that does not result in a change to the DNA sequence itself. Information can be heritable through cell division and/or organismal generations. Epigenetics has a strong association with a variety of diseases; identification of the perturbations of epigenetic marks may therefore act as a biomarker for early disease identification.

Individuals in a population vary not only in terms of DNA but also in terms of epigenetics. Thus, their response to treatment or their disease progression can be dependent on the epigenetic modifications they possess. For this reason, individual variation in terms of epigenetics may affect patient outcome and so it is important to be able to identify patients who may respond more or less favorably to specific treatments because of their epigenetic status. Epigenetic modifications are generally stable, but they can also be reversed, leading to the potential for personalized treatment to reverse undesirable epigenetic states or to introduce beneficial ones.

This chapter focuses on the influence of DNA methylation, histone variants, and histone modifications as sources of epigenetic information. (For conciseness, the chapter does not cover other epigenetic mechanisms such as noncoding RNA/microRNAs, chromosome structural conformation, prions, or RNA feedback loops.) Methods of measuring and manipulating epigenetic marks are discussed. A summary of the known effects of and disease relationships to epigenetics and prospects for epigenetic therapies are provided.

4.2 DNA METHYLATION

DNA methylation is the addition of a methyl group to cytosine (C) or adenine (A) DNA nucleotides. Most organisms are capable of methylating cytosine bases (methylcytosine), which leads to methylation being frequently referred to as the "fifth base"; however, methyladenine is generally regarded as a bacterial DNA feature. Methylation is a relatively stable and heritable modification of DNA nucleotides that serves to regulate a number

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00004-9
© 2014 Elsevier Inc. All rights reserved.

of cellular functions—for example, during cellular differentiation. Although, in specific cases, DNA sequence can vary between cell types (i.e., immune class switch recombination), in the majority of cases DNA is identical across all of an organism's cells. In contrast, DNA methylation frequently varies between different host organism cell types and, although relatively stable, is potentially reversible.

Because the focus of this book is on pharmacogenetics, this chapter concentrates on the methylation of cytosine nucleotides that occurs in mammalian genomes, with an emphasis on human molecular biology.

Methylation occurs predominately at CpG dinucleotides in mammalian genomes, where a CpG site is a cytosine immediately followed by a guanine when reading 5′ to 3′ on a linear piece of DNA—the "p" represents the single phosphate linking the two nucleotides. Because a CpG is a palindromic sequence, it has a complementary pair on the opposing strand.

CpG sites in human genomes are heavily methylated—on average, 70% of CpGs in primary human cells. Cytosines in any other configuration (CHG and CHH sites, where "H" refers to any base other than guanide) are methylated much less frequently (~0.1%). A methylated CpG (meCpG) is selected against in the genome because of hypermutability resulting from spontaneous deamination of the cytosine. While deamination usually forms a uracil—a nucleotide not normally found in DNA that can be detected and replaced by DNA repair enzymes—the deamination of methylated cytosines forms a thymine, which is commonly found in DNA and is less easily detected. Nonetheless, Thymine-DNA glycosylase (TDG) is an enzyme that attempts to repair such CpG mutations, but it is not fully effective.

Consequently, CpG dinucleotides occur at a rate of only about 1% in the genome; this is around one-quarter of what would be expected by chance. Surprisingly, however, there are concentrations of CpG dinucleotides in the genome, known as CpG islands. These tend to be located near gene transcription start sites (particularly at housekeeping genes), in enhancers, and occasionally in genes themselves. The cut-off for defining a CpG island usually includes a minimum size, a minimum percentage of G+C nucleotides, and a minimum ratio of observed-to-expected CpG content compared to total G+C content. Cut-offs for GC content and Obs/Exp ratio are often based on historical values of 50% and 0.6, respectively, with a minimum size of 200 bp [1].

In normal human cells, CpG islands are generally unmethylated—in stark contrast to the 70% methylation seen genome-wide. As a result, they are usually explained as the result of selective pressure for high CpG content in a region and/or the result of CpGs being unmethyated for another reason, such as gene regulation.

4.2.1 Effects of DNA Methylation

DNA methylation has been implicated in the regulation of gene expression; however, the exact mechanism by which this process occurs is still poorly understood. The methylation status of CpG islands located near the transcription start site of a gene has been shown to be linked to the expression level of the gene, and it is believed that lack of methylation at CpG islands permits expression while increased methylation limits expression to low levels. Although the vast majority of CpG islands in normal cells are constitutively unmethylated [2], a number become methylated in some cases, and this methylation is associated with transcriptional silencing [3]. Focal methylation and silencing of some tumor suppressor genes [4] have also been observed in cancers [5]. However, methylation may be a secondary effect that occurs after gene silencing [6] rather than the causative agent. Similarly, loss of methylation at oncogenes has been associated with reactivation in tumors [7].

Like the genome-wide tendency toward high levels of CpG methylation, gene bodies tend to be highly methylated as well. The relationship between gene expression and gene body methylation is difficult to decipher because, while some cell lines exhibit a positive correlation between gene expression and gene body methylation, this relationship has been shown to be nonmonotonic [8]. Specifically, genes with the lowest and highest levels of expression have been shown to have the lowest level of methylation, while genes with intermediate levels of expression have been shown to possess the highest level of methylation in the gene body.

One proposed role for the methylation of gene bodies is the suppression of spurious transcription initiation. A negative correlation exists between levels of methylation in gene bodies and the initiation of intragenic transcripts at the SHANK3 locus [9]. It is unclear if this correlation is a mechanistic result of the methylation or if the methylation and the suppression of spurious intragenic transcription are both the result of a common chromatin structure [10].

Methylation has been associated with the suppression of DNA repeat transcription and retrotransposition. Transposable elements [11] are DNA sequences that can duplicate themselves in the genome via an RNA intermediary (i.e., class I elements) or change their position within the genome via transposase enzymes (i.e., class II elements). Transposable elements are potential mutagens and can damage the genome by insertion into functional genes or by leaving unrepaired gaps from the location that the element left [12]. Methylation has been shown to suppress transcription of endogenous retroviruses [13], and methylation of retrotransposons may be necessary for mammalian viability [14].

Although methylation of specific CpG sites is capable of suppressing transcription of LINE transposable elements [15], demethylation of the same elements is not

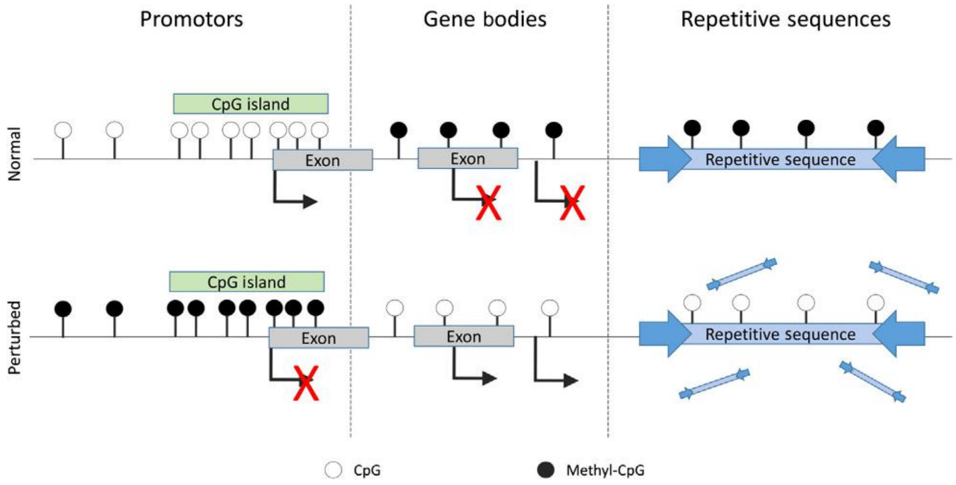


FIGURE 4.1 **DNA methylation can occur in different locations throughout the genome.** A variation from normal behavior is associated with disease. In the normal case, CpG islands are unmethylated while gene bodies and repeat regions tend to be methylated. This allows normal transcription originating from the transcription start site of the gene and transcription suppression elsewhere. In the perturbed case, CpG islands gain methylation while the gene body loses it. This inhibits transcription originating from the transcription start site while allowing spurious transcription in the gene body. Repeat regions also lose methylation; this has been linked to activation of repetitive elements, recombination, and genome instability.

always sufficient on its own to reactivate transcription [16]. Numerous cancers have been observed to have large-scale loss of methylation compared to matched normal control tissues [5,17,18], and it is thought that this hypomethylation may be responsible for reactivation of transposon elements, leading to decreased genome stability and the development of further genomic mutations in the tumor as shown in Figure 4.1.

Methylation has a well-documented effect on allele-specific expression. Normal development requires the contribution of both maternal and paternal genomes, but some genes are expressed only from the alleles of either the maternal or the paternal source. Artificially created genomes where both alleles have maternal (gynogenone) origins have double the expression of maternally derived genes and zero expression of paternally derived genes. The reverse is true of genomes with two paternal contributions (androgenones). These artificially generated genomes have been studied to determine that many allele-specific genes are involved in embryonic development.

One particular example of genomic imprinting is the IGF2r gene (insulin-like growth factor 2 receptor), which depends on the methylation of an intronic CpG island to control expression [19]. In the maternal copy of the gene, the promotor is unmethylated and active, which allows the production of the IGF2r mRNA, while an intronic CpG island is methylated. This intronic CpG island acts as a downstream promotor that is capable of producing an antisense RNA that can interfere with the gene's regular functioning. In the maternal copy of the gene, the intronic CpG island is inactivated by methylation. In the paternal copy of the gene, on the other hand, the intronic CpG island is unmethylated while the promotor is methylated. This allows the production of the antisense RNA via the downstream promotor coupled with transcription suppression from the coding strand via methylation of the gene promotor.

Several imprinted genes have been associated with cancers either as oncogenes (e.g., H19) or as tumor suppressors (e.g., CDKN1C). It is thought that, at least in some cases, methylation controls expression of imprinted genes via the modulation of enhancer access (via the CTCF protein) to the promotor [20]. In this model, CTCF suppresses expression of the IGF2 gene (insulin-like growth factor 2) via blocking of enhancer elements. This imprinting is then made allele-dependent by differential methylation patterns. Hence, DNA methylation controls gene expression by binding insulator proteins that can deny enhancer proteins access to the gene.

Another example of DNA methylation–linked silencing is chromosome imprinting (X-inactivation). In female mammalian genomes, one copy of chromosome X is largely inactivated by the formation of heterochromatin—although this does not affect every gene on the chromosome. Heterochromatin formation acts as a dosage compensation mechanism and ensures that the expression of genes located on chromosome X is roughly identical between males (one X chromosome) and females (two X chromosomes). Methylation is thought to promote the silencing of genes on the inactive X chromosome, although it occurs after gene repression and so may act to stabilize silencing. The inactive X chromosome is highly methylated at transcription start sites, but seemingly becomes hypomethylated across gene bodies and intergenic regions when it is inactivated [21]. Methylation therefore serves to imprint and reinforce an absence of transcription in the inactive X chromosome [22].

4.2.2 DNA Methyltransferases

Methylation is deposited and maintained by a family of catalytic enzymes known as DNA methyltransferases (DNA MTases, or DNMTs). These are categorized into two distinct classes: maintenance methylation and de novo methylation.

In normal somatic human cells, maintenance methylation occurs during every round of cell division. During DNA replication it is necessary to duplicate the original methylation patterns onto the newly synthesized daughter DNA strands. If maintenance methylation is interrupted or suppressed, the daughter strand of does not become methylated; this subsequently leads to passive demethylation. Maintenance methylation is primarily carried out by DNMT1. DNMT1 has a preference for hemimethylated CpG di-nucleotides (CpG di-nucleotides with discordant DNA methylation between the two strands), which are produced during replication [23] and relocated at replication foci. Mice that have heterozygous mutations (i.e., one working copy) in DNMT1 are indistinguishable from wild-type mice; however, homozygous mutants are embryonically lethal [24].

Unlike embryos, embryonic stem cells grown in culture with homozygous mutations are viable [24] with low but stable levels of methylation [25] that can be maintained for over 20 cell generations. This was one of the original pieces of evidence for the existence of methyltransferases which are independent of both DNMT1 activity and the cell cycle. DNMT3a and DNMT3b were identified through their protein sequence similarity to bacterial methyltransferases, and were shown to have equal catalytic activity on both unmethylated and hemimethylated CpGs [26], such that they are capable of de novo DNA methylation.

The DNMT3 family does not appear at replication foci [27], unlike DNMT1; instead, it is expressed throughout the cell cycle. The de novo methyltransferases are essential in establishing the methylation pattern of the organism and in de novo methylation during development [28] and embryogenesis [29]. Analysis indicates that DNMT1 may also have a role to play in methylation outside of replication. It is kinetically very slow in comparison to the DNA replication fork cycle [30], and it is difficult to see how it can keep up with the replication fork. DNMT1 has also been shown to be active during the G2 (cell growth and protein synthesis) and M (mitosis) phases of the cell [31], and activity outside of the S phase (DNA synthesis) is dependent on the TS (targeting sequence) domain in DNMT1 rather than on the PBD (PCNA-binding domain) that is used at the replication fork. Furthermore, DNMT1 interacts with the DNMT3 family [32], possibly as part of de novo methylation activity.

It is worth mentioning a protein previously classified as DNMT2 that has strong sequence similarity to the other methyltransferases and was once thought to be another DNA methyltransferase of unknown biological function. Ultimately, DNMT2 was found to methylate aspartic acid in transfer RNAs (tRNAs), thus protecting them against ribonuclease cleavage [33]. It was renamed TRDMT1 to reflect the change in understanding of its function.

A further modification that can be carried out on already methylated cytosine bases is the addition of a hydroxy group, which forms 5-hydroxymethylcytosine (5hmC). Hydroxy groups can be added by the oxidation of methylated cytosines by the Tet family of enzymes, and may be followed by decarboxylation in order to perform active cytosine demethylation [34]. Other functions of 5hmC are still unclear, although it is suspected to play a role in imprint erasure in germline cells [35], has been implicated in demethlyation of LINE-1 repeat elements [36], and is abundant in the brain [37], suggesting a role in regulation of neuron function.

4.2.3 Techniques for Measuring Methylation

There are a number of techniques for measuring DNA methylation levels—both for investigating specific sites or loci and for genome-wide approaches.

One cell-based approach is to stain the cells with an antibody that binds to 5-meC; this can be visualized using immunofluorescence or immunohistochemistry to obtain a gross overview of the abundance of methylation in the genome. Such an approach is frequently used to quantitate global levels of methylation in cells—for example, at different stages of development [38]—by observation of changes in the global level of methylation or changes in localization. However, because resolution is poor this technique can be used only to measure gross methylation levels.

5-meC antibodies can be further used to immunoprecipitate methylated DNA sequences (MeDIP). Genomic DNA is extracted from cells and sonicated into small fragments. It is incubated with 5meC antibodies and immunoprecipitation is used to remove unbound DNA. The remaining DNA is purified away from the antibody [39]. The quantity of DNA recovered from different loci can subsequently be measured by polymerase chain reaction for a single locus (MeDIP-PCR), by microarray for many preselected loci (MeDIP-chip), or by next-generation sequencing for a genome-wide view (MeDIP-seq). The resolution of MeDIP is limited by the size of the DNA fragments after sonication. Smaller fragments give greater resolution for PCR and array-based approaches, but they do not bind as efficiently with the 5meC antibodies (which bind better with multiple methylated sites) and pose targeting problems in designing unique primers (PCR) and probes (microarray). If fragments are very small, they can inhibit unique alignment of the sequenced fragment (NGS sequencing), which means that the resolution of MeDIP is practically limited to around 300 bp. The binding efficiency of the antibody, as well as the subsequent amount of DNA that is precipitated, is a function of both the methylation level of the fragments and the density of CpG di-nucleotides in them; thus, statistical tools are used to estimate the absolute methylation level by accounting for a region's CpG density [40].

A recent innovation is the use of affinity purification via the naturally occurring methyl-binding domain (MBD), extracted from the human MBD2 protein, which binds with symmetric meCpGs in double-stranded DNA. DNA is then captured through elution in steps, using a salt gradient that allows affinity-based capture [41] to achieve greater sensitivity than the comparable antibody-based approach along with greater yields. Importantly, the MBD approach utilizes a more flexible protocol and allows fractionation into low-, intermediate-, and high-CpG density fragments via variation of the elution buffer's salt content.

It is possible to obtain higher-resolution data on the distribution of methylation at individual CpGs through treatment of DNA with bisulphite ($NaHSO_3$), also known as sodium bisulfite. The addition of bisulphite to DNA triggers the conversion of cytosine to uracil, except where the cytosine is protected by a methyl group. The genomic loci of interest can then be amplified by PCR (which subsequently converts uracil to thymine) and subsequently directly sequenced [42]. Methylation-specific PCR, a recent development, removes the need to actually sequence the locus of interest [43] by designing complementary pairs of PCR primers that are methylation-specific and unmethylation-specific and are used in parallel. Methylation level is determined by the respective ability of the primers to anneal specifically to the region of interest and therefore their ability to amplify the DNA region under examination.

Microarrays were developed as an extension of the techniques just described. DNA is bisulphite-converted and hybridized to a microarray that contains between hundreds and hundreds of thousands of probes for individual CpG di-nucleotides. Each probe consists of a pair of oligonucleotides specific to the CpG targeted; one probe contains a CGN sequence and the other contains a TGN sequence, so representing methylated and unmethylated CpG status respectively [44]. The level of methylation for a given probe can then be calculated by the ratio of florescence between the two oligos. Modern arrays range from 27,000 to 450,000 probes [45].

With the development of next-generation sequencing (NGS) technology, it has become possible to directly sequence the bisulphite-converted DNA fragments at high coverage. Genome-wide bisulphite sequencing allows quantitation of virtually all CpG di-nucleotides in the genome as well as examination of cytosines in non-CpG contexts. This has allowed the examination of large domains of methylation [46], which previously had not been readily accessible to approaches forced to target focal regions of interest—typically, CpG islands and genic regions.

As NGS is extremely expensive, it is still necessary to consider more cost-effective approaches to collecting methylation data for the whole genome. Reduced-representation bisulphite sequencing (RRBS) is one such approach that first uses a restriction enzyme (Mspl) to cut genomic DNA at CCGG sites [47]. Size-selecting the resulting fragments to the 40–220 bp range results in selection toward most CpG islands in the genome. It is then possible to bisulphite-treat this DNA as normal and sequence it using NGS technology to obtain high-coverage, single-base-pair-resolution methylation data for most of those CpG islands. Similarly, it is possible to specifically target one or more extended genomic regions of interest by using targeted enrichment approaches [48] to enrich specific regions of DNA—for example, by using a library of thousands or millions of baits to hybridize to and select for targeted DNA sequences. Bisulphite sequencing cannot differentiate between 5mC and 5hmC because both types prevent the conversion of cytosine to uracil. One modification for differentiating the two is two-step oxidative bisulphite sequencing (OxBS-seq) [49]. Oxidation of 5hmC to 5fC (5-formylcytosine) is first achieved via potassium perruthenate, which does not modify either unmethylated or 5mC methylated cytosines. Bisulphite treatment is then carried out to convert both unmethylated and 5fC cytosines into uracil, and in this way the OxBS-seq can specifically record only 5mC methylation. By performing traditional and oxidative bisulphite sequencing in parallel, it is possible to subtract the 5mC-specific readout from the total methylated bisulphite signal to determine 5hmC distribution.

Another modified approach is Tet-assisted bisulphite sequencing (TAB-seq), which uses the enzyme TET1 to convert 5mC into 5caC (5-carboxylcytosine) [50]. First a glucose is added to 5hmC using β-glucosyltransferase (βGT) to form 5gmC (β-glucosyl-5-hydroxymethylcytosine). This protects the existing 5hmC from TET-mediated oxidation. After all existing 5hmC is blocked, all 5mC is converted to 5caC using the TET1 protein. Bisulphite treatment of the resulting DNA converts all unmethylated cytosines to uracil and all 5caC (originally 5mC) to 5caU, while the blocked 5hmC remains protected as 5gmC. Sequencing results in a cytosine base originally 5hmC and a thymine base originally unmethylated or 5mC-methylated. This provides specificity in determining the levels of 5hmC by excluding 5mC methylation.

4.3 HISTONES AND VARIANTS

A nucleosome is the basic structural unit in which DNA is packaged. A single DNA base pair is approximately 0.34 nm long [51]. The human genome contains 6 billion such bases (3 billion–base diploid genome), which is approximately 2.04 m DNA per cell. Histones are highly conserved basic proteins that form a nucleosome by wrapping 147 base pairs

TABLE 4.1 Known Human Histone Variants Obtained from the UniProtKB Protein Database

Histone Family	Variants
H1	H1.0, H1oo, H1x, H1.1, H1.2, H1.3, H1.4, H1.5, H1t
H2A	Macro-H2A.1, macro-H2A.2, H2A.J, H2A.V, H2A.x, H2A.Z, H2A type 1 (A,B/E,C,D,H,J), H2A-Bbd type 1, H2A type 2 (A,B,C), H2A-Bbd type 2/3, H2A type 3
H2B	H2B type 1 (A,B,C/E/F/G/I,D,H,J,K,L,M,N,O), H2B type 2 (C,D,E,F), H2B type 3-B, H2B type F (M,S), H2B type W-T
H3	H3-like centromeric protein A, H3.1, H3.1t, H3.2, H3.3, H3.3C
H4	No known sequence variants

of DNA around two copies of each of the four core histones (H2A, H2B, H3, and H4) in an octamer. The very basic nature of histones allows a tight binding with the acidic DNA to form a characteristic "beads on a string" structure.

The H1 linker histone does not form part of the core nucleosome, but instead can bind at the nucleosome's DNA entry/exit point to lock the DNA into place and, it is thought, to allow nucleosomes to be pulled together to form higher-order structures. Unlike the four core histones, there is only one copy of the H1 linker histone per nucleosome.

The combination of nuclear DNA, RNA, and proteins (including histones) is referred to as the chromatin. Open chromatin, or euchromatin, refers to chromatin considered to be easily accessible, lightly packed, and often transcribed in interphase cells. In contrast, closed chromatin, or heterochromatin, remains condensed in interphase and is thought of as tightly packed to prevent transcriptional activity. Heterochromatin is often subdivided into constitutive heterochromatin, located at highly repetitive regions (e.g., pericentromeres and telomeres) and always present, and facultative heterochromatin, which can be located anywhere in the genome and often performs a regulated gene-silencing function (e.g., the inactive C chromosome [52]).

While histones are among the slowest-evolving proteins, there are some variants that result in sequence differences. Examples of these variants are listed in Table 4.1. Histone variant names are based on UniProtKB protein names. Several histone variants have multiple subtypes, which are indicated in parentheses and separated by commas in the table. Some subtype names include a forward slash, indicating that the variant is transcribed from multiple genes but has the same amino acid sequence. Histone variants, in general, may be transcribed from multiple genes (data not listed) but generate the same amino acid sequence.

The chromatin structure can be modified by different histone variants to create different chromatin environments in localized portions of the genome, and sometimes in a tissue-specific fashion (e.g., testes variants). This allows chromatin to regulate a number of DNA functions such as gene expression and DNA damage [53].

Histone H1 has a number of variants that range from 21–29 kDa (194–255 amino acids). Its variants are often classified by the strength of their linker function into strong (H1.0, H1.4, H1.5, H1x), intermediate (H1.3), and weak (H1.1, H1.2) based on their ability to condense the chromatin [54]. Histone H1 can be further categorized by tissue origin; H1.2 and H1.4 are found in most somatic cell nuclei while some H1 variants (H1t, H1oo) are germline-specific and regulate chromatin structure during spermatogenesis and fertilization [55,56].

Histone H2A has the largest number of variants, most of which are approximately 14 kDa (~130 amino acids) although some vary considerably in size. Some H2A variants appear to have been highly conserved through evolution (i.e., H2A.Z) and have been associated with active or potentially active chromatin [57], particularly around the transcription start site of active genes [58]. The macroH2A variant are much larger version of H2A at around ~40 kDa (327 amino acids). They mark repressed chromatin in a signal-dependent fashion [59] and is enriched on the inactive X chromosome [60], where it is thought to suppress transcription in part through inhibition of PARP-1 [61]. In contrast, the H2A Barr body-deficient (H2A-Bbd) variant wraps only 118 base pairs of DNA and is found only on the active X chromosome [62], where it is associated with transcriptionally active regions of the genome.

Histone H2B is another 14 kDa protein, consisting of approximately 126 amino acids and having several variants, although their roles are not well understood. One of the few variants with a known function is H2B W-T, a testes-specific variant; nucleosomes containing it have been shown, unlike conventional H2B, to be unable to recruit chromosome condensation factors [63].

Histone H3 is 15–16 kDa protein (~135 amino acids). Its three main forms are H3.1, H3.2, and H3.3. H3.1 and H3.2 differ by only one amino acid, although they may serve different roles. H3.1 has been shown to be localized to regions that can be switched to active or repressed based on the state of the cell, while H3.2 is thought to be enriched in transcriptionally repressive domains [64,65]. In contrast, H3.3 is often associated with transcriptionally active chromatin and active and poised gene promoters and regulatory elements, and is thought to replace existing H3

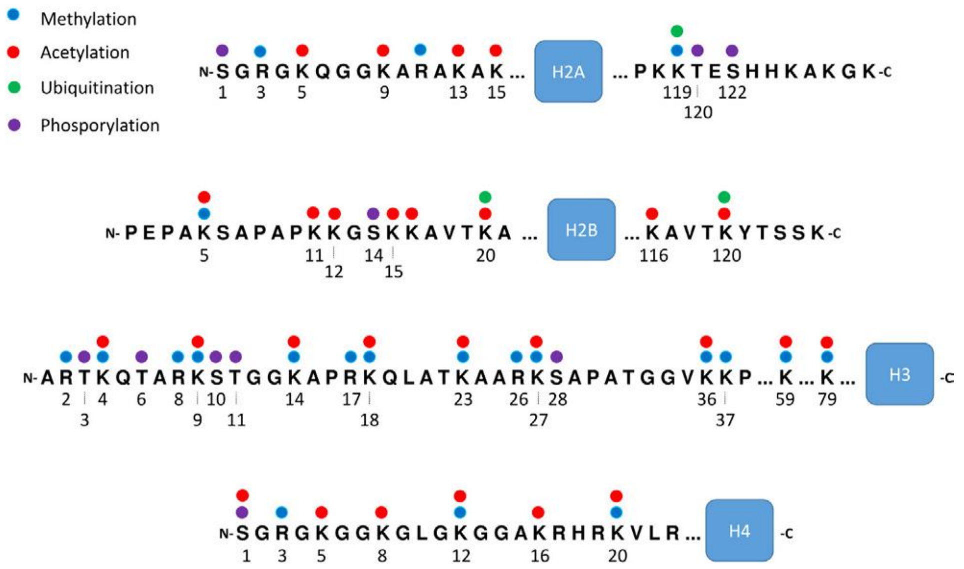


FIGURE 4.2 **Histone modifications are prevelant on the histone tails.** Shown is a representative sample of the most commonly studied histone modifications (methylation, acetylation, ubiquitination, and phosphorylation) for the four core histones. Modifications can also affect the globular region of the histone (not shown). Numbers indicate amino acid position in the histone sequence.

histones during transcription [66]. H3.3 is also recruited to telomeres during telomere replication and processing [67] in stem cells. CENPA is a 17 kDa H3-related protein that drives the assembly of centromeres [68].

H4 is 11.3 kDa (102 amino acids) and is the most conserved histone; it has no described human sequence variants [69]. This is possibly because it has docking sites for all three of the other core histones [70] and any significant variation may compromise the structure or stability of the nucleosome.

The bulk histone content in a cell is duplicated when the cell passes through the S phase (DNA synthesis phase) of the cell cycle. This relatively short phase results in an extremely high demand for histones to create the chromatin on the newly synthesized DNA. Histone transcripts created during S phase are termed replication-dependent and are characterized by their lack of introns and polyadenylated 3′ end; instead, they have a purine-rich sequence forming a stem-loop structure [71]. SLBP (stem-loop–binding protein) binds to this sequence and facilitates mRNA stability, nuclear export, and translation in proliferating cells [72]. In contrast, replication-independent histones (such as H1.0 and H3.3) are transcribed throughout the life of the cell and feature a poly(A) tail to facilitate canonical mRNA metabolism via poly(A)-binding proteins (PABPs).

Histones can be added and removed from DNA by chaperone complexes that allow the histones in each nucleosome to be changed according to gene expression programs. Examples of histone chaperones include HIRA (replication-independent) and DAXX (replication-dependent), which are both responsible for deposition of histone H3.3 to DNA.

4.3.1 Histone PTMs

In addition to the numerous histone variants, there are a multitude of post-translational modifications (PTM) that can change the behavior and functionality of the histone-containing nucleosome. Each of the core histones contains an N-terminal tail protruding from the nucleosome on which most of these modifications take place. Amino acids in the tails can be subjected to many covalent modifications, including methylation (mono-, di-, tri-), acetylation, phosphorylation, and ubiquitination. Other modifications include sumoylation, citrullination, and biotinylation. Modification of the histone tail can manipulate chromatin formation directly or may allow or disallow the binding of different transcription factors. Some modifications are mutually exclusive (i.e., they occur at the same amino acid), while others may be combinatorial [73], occurring simultaneously on different residues to regulate chromatin function. A summary of commonly studied histone post-translational modifications is provided in Figure 4.2.

Post-translational modifications are identified by a standard nomenclature indicating (1) name of the core histone; (2) amino acid affected using the IUPAC single-letter codes (e.g., K for lysine, R for arginine); (3) amino acid position in the protein (starting from the N-terminus); (4) type of modification (me: methyl, ph: phosphate, ac: acetyl, ub: ubiquitin); and, optionally, (5) number of modifications (one, two, or three). For example, H3K4me3 indicates tri-methylation of Lysine 4 on histone H3.

Histone modifications are carried out by families of enzymes, such as histone methyltransferases (HMT) or acetyltransferase (HAT). Similarly, they can be reversed by histone demethyltransferases (HDMT) or Deacetylase (HDAC) enzymes.

4.3.2 Histone PTM Effects

There are many potential consequences of post-translational modification of histones. Trimethylation of lysine 4 on histone H3 is associated with activated transcription when located near the transcription start site to activate genes [74]. H3K36me3 in the gene body similarly indicates actively transcribed genes. However, methylation of lysines 9 and 27 typically correlates with transcriptional repression [75].

Histone acetylation removes the positive charge, which relaxes the interaction with negatively charged DNA and is thought to make it more euchromatic and accessible to transcription [76,77], DNA repair [78], and histone deposition [79]. Post-translational modifications have been associated with numerous other biological functions and contexts, including polycomb-mediated prevention of transcription elongation during spermatogenesis [80], apoptosis [81], and mitosis [82].

Each histone can be covalently modified, and each can be modified on multiple amino acids. Several amino acids support multiple types of modification, which has raised the notion of a parallel histone code or histone language [83] that is perhaps as important as the DNA code. The combinatorial effects of histones are significant. One such effect is the ability of histone modifications to cooperate and reinforce each other. For example, inactive regions/heterochromatin are marked by low levels of acetylation, low levels of activating marks (H3K4me, H3K36), and high levels of inactivating marks (H3K9me, H3K27me) while in active regions/euchromatin this trend is reversed [73,84].

In addition to individual effects of single modifications or the cumulative effect of modifications with similar functions, combinations of histone modifications can encode more dynamic behavior. One typical example is bivalent domains [85] found in embryonic stem cells where genes are simultaneously marked with both repressive (H3K27me3) and activating (H3K4me) histone modifications. The repressive mark takes precedence over the activating mark. However, the gene is kept poised and can be activated by simple removal of the repressive mark. This may allow more rapid activation of the gene than would be possible with the sequential removal of H3K27me and the subsequent addition of H3K4me.

Similarly, histone modifications may interact with each other even where they do not share the same amino acid. H3K36me3, while associated with active transcription, is known to recruit a HDAC complex [86] that prevents the acetylation of histones and the subsequent initiation of transcription in the gene body.

4.3.3 Techniques for Measuring Histones and Other Chromatin-Interacting Proteins

A frequently used approach for measuring protein interaction with DNA is chromatin immunoprecipitation (ChIP) [87]. DNA is first cross-linked to form a covalent bond between physically nearby proteins and DNA (i.e., both protein–DNA and protein–protein bonds). As in MeDIP, an antibody is used to precipitate the protein of interest and the DNA fragments associated with the protein are then purified. ChIP can be combined with PCR (ChIP-PCR) to probe a specific region of interest; it can be used with a tiling array (ChIP-chip); or it can be combined with NGS technologies to determine genome wide protein–DNA binding (ChIP-seq). Like MeDIP, the resolution of ChIP is limited to a few hundred base pairs; however, it is possible to use exonucleases to degrade exposed DNA before sequencing to improve resolution [88], although this comes at the cost of shorter reads from the DNA sequencer that are subsequently more difficult to uniquely map to the genome.

One difficulty with ChIP is that it relies on the availability of specific and sensitive antibodies to each protein of interest, whereas MeDIP relies only on a single well-characterized antibody. In the case of histone variants, or post-translational modifications, specific antibodies may not be readily available because of the small differences between variants and PTMs. Thus, antibodies may not be able to specifically recognize a given histone variant/PTM or may recognize multiple variants/PTMs. It is therefore vitally important that ChIP studies for histones employ extensive validation and screening for antibodies, using peptide arrays to check for antibody specificity and sensitivity.

An alternative to the ChIP procedure is labeling the DNA close to the binding site of the protein of interest. DamID [89] involves the creation of a fusion protein composed of Dam (DNA adenine methyltransferase), which has been extracted from E-coli and methylates adenine, and a DNA-binding protein (i.e., the protein of interest). Expression of this fusion protein leads to the methylation of adenine in GATC sequences near the protein's binding site without any other noticeable effects. Methylated adenine is absent in eukaryotes, so any methylation of GATC sequences is likely due to the proximal binding of the protein of interest. DNA is then digested with the DpnI enzyme, which cuts only methylated GATC sites. The sequence at the ends of the DNA fragments is known, so adapters can be ligated to the ends resulting from digestion and so PCR with matching primers can be carried out. The result is that fragments flanked by the methylated GATC sites are amplified.

Alternatively, a methylated adenine-specific antibody can be used, as in MeDIP. The Dam protein itself has intrinsic DNA-binding activity, so it is necessary to compare the Dam–protein fusion to a Dam-only negative control to filter out "background" adenine methylation as a result of Dam binding. DamID can be combined with microarray technologies (DamID-chip) [90] or with NGS technologies now available [91].

A modification of the bisulphite-sequencing approach can measure nucleosome occupancy simultaneously with assessment of DNA methylation (NOMe-Seq). Briefly, this involves the addition of a GpC methyltransferase (M.CviPI), and it relies on differences in methyltransferase accessibility to GpC sites in isolated chromatin. Because the majority of GpC sites are physiologically unmethylated (only possible in a GpCpG context), it is inferred that, upon sequencing, any GpC sites that are methylated are the result of exogenous M.CviPI. If a GpC site is unmethylated, it is further inferred that this is because of the enzyme's inaccessibility and the site is subsequently used as a readout of nucleosome occupancy [92]. Simultaneously, *in vivo* methylation status can be read from the unaltered CpG sites and, importantly, can be retrieved from the same molecule—allowing a direct comparison of methylation and nucleosome occupancy in a single cell. Unfortunately, mC in a GpCpG context is difficult to interpret by this method.

One final approach is the combination of ChIP with bisulphite sequencing. The DNA fragments that are obtained from ChIP can be bisulphite-treated such that the methylation status of the protein's binding domain can be examined—a process referred to as ChIP-BS [93]. ChIP-BS allows examination of methylation in the context of specific transcription factors and/or histone modifications—for example, to determine if methylation status and a particular chromatin-binding protein are related.

4.4 EPIGENETIC DISORDERS AND THERAPIES

There are a number of imprinting-related disorders that occur either from heterozygous loss of an imprinted gene or from loss of imprinting. Chromosome 15 q11–13 is associated with the development of Prader-Willi syndrome (PWS) and Angelman syndrome (AS) [94]. This region contains genes that are expressed only from maternal or paternal contributions. Mutation, deletion, or loss of function of a single gene (UBE3A), which is expressed exclusively from the maternal allele, is sufficient to cause AS despite the presence of an intact but unexpressed copy on the paternal allele. Loss of paternal contribution of this region leads to PWS, although the specific genes involved are less well established—many genes in the region are expressed only from the paternal contribution, but none have been linked definitively to the condition. The most likely candidate is a snoRNA gene cluster [95,96], but this does not seem sufficient on its own to cause the disorder [97]. PWS may therefore be the result of the combinatorial loss of imprinted genes in the snoRNA cluster and other, as yet unknown gene abnormalities.

In addition to imprinting disorders caused by loss of an individual parental genetic contribution, loss of imprinting itself can be due to the disruption of epigenetic marks (gain or loss of methylation). The IGF2/H19 locus is well studied in terms of imprinting disorders and is known to be deregulated in cancers [98] and in other disorders such as Beckwith-Wiedemann syndrome [99]. The protein CTCF is involved in regulating expression of this gene. It is sensitive to methylation status at its binding site [100] and acts as a chromatin insulator that divides the genome into functional domains; it also regulates the ability of enhancers and transcription factors to access promotors and prevents increases in methylation [101]. Loss of imprinting at the IGF2/H19 locus is associated with loss of CTCF because of methylation changes, without any apparent mutations of CTCF or its binding site [102].

Many cancers have large domains of DNA hypomethylation [18,103,104] along with more focal areas of hypermethylation [5,17,105]. Hypomethylation is an early event [106] and increases with disease stage [16], suggesting that it has roles in both cancer initiation and maintenance. Hypermethylation of CpG islands is thought to play a role in repressing tumor suppressor genes—silencing of these genes provides a growth advantage for the affected population of cells and increases their survivability [107]. Treatment with 5-azacytidine and 5-aza-2′-deoxycytidine, which block maintenance methylation during replication, has been proposed as cancer therapy [108,109] to inhibit CpG island hypermethylation; both are approved for use in humans in several countries for cancers such as AML and MDS in several countries.

Huntington's disease is caused by a mutant huntingtin gene (Htt) that in normal cases has a CAG (glutamine) repeated fewer than 26 times. Huntington's disease is caused by the extension of this repeat in at least one allele of the gene to a sequence of greater than 36 repeats, resulting in a protein that causes an increased decay rate of neurons [110]. One hypothesis is that mutant huntingtin binds to CBP and/or P/CAF and reduces their availability to mediate histone acetylation at genes that promote neuronal survival. Treatment with HDAC inhibitors (to compensate for loss of CBP and/or P/CAF function) has been shown to help in drosophilia [111] and mouse [112] models of Huntington's. HDAC inhibitors are also considered a suitable approach for restoration of learning and memory in neurodegenerative diseases [113] and as anti-cancer therapies [114], although their method of action is still not completely known. HDAC inhibitors have also been approved for human use in some locations.

Histone variants and modifications have been linked to aging. Specifically, H4K16 acetylation's regulation of cellular lifespan [115] is known to be through interaction with Sir2. Sirtuins, including Sir2, are a family of conserved deacetylases with connections to longevity and/or healthspan in a number of species. In yeast, Sir2 deacetylates H4K16ac; an age-associated decrease in expression of Sir2 correlates with an increase in H4K16ac and impaired transcriptional silencing. Suppressing H4K16ac levels via deletion of an acetyltransferase (Sas2) increases cell lifespan. Mutations in H4K16 and H3K56 have been shown to result in reduced lifespan [115]. Furthermore, methylation of histone 3 is known to affect the lifespan of worms via H3K4me3 in the germline [116] and in somatic cells via H3K27me3 [117].

Histone modifications have been further implicated in cancer, and there is evidence that cancers can be promoted by abnormalities in histone methylation. Global loss of H4K9me3 has been observed in multiple cancer lines [118], while hypermethylation of H3K9 and hypomethylation of H3K4 are characteristic of transcriptionally silenced heterochromatin in cancers. Histone demethylase (HDM) activity has been further linked to neurological disorders and cancer [119].

There is evidence that the level of H3K27 methylation is important for normal cell growth and that deregulation of this modification is associated with multiple cancers. H3K27 methylation is catalyzed by EZH2—the histone methyltransferase that modifies H3K27—and mutations or disregulation of this gene are found in multiple cancers [120,121]. High EZH2 expression is prognostic of poor patient progression and for this reason EZH2 inhibitors are an active area of research [122].

4.5 CONCLUSION

This chapter summarized the mechanisms and effects of DNA methylation, histone variants, and histone modifications. Tools for measuring and manipulating each of these epigenetic marks were presented. A brief summary of the role of epigenetic marks in diseases such as cancer was also provided.

While methylation and histone variants/modifications have individual effects, they have also been shown to interact at a molecular level, so it is important to consider possible combinatorial effects. DNA and histone methylation have been shown to cooperate, and DNA methylation is essential to ensure that histone methylation (H3K9me) is maintained through DNA replication, mediated by the Methyl-CpG–binding protein (MBD1) [123]. H3K9me is required for DNA methylation in neurospora crassa [124] and arabidopsis thaliana [125].

The study of epigenetics is important for understanding cell function, differentiation, and behavior in disease. Because so many diseases involve changes to methylation and/or histone variants and modifications, it is important to understand the effect they have so that they can be reversed with clinical treatments or used as biomarkers for diseased cells. As these changes are reversible (unlike DNA mutations), the process of manipulating them should be more feasible through drug-based treatments.

REFERENCES

[1] Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol 1987;196(2):261–82.
[2] Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 2008;454(7205):766–70.
[3] Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 2008;9(6):465–76.
[4] Merlo A, Herman JG, Mao L, Lee DJ, Gabrielson E, Burger PC, et al. 5' CpG island methylation is associated with transcriptional silencing of the tumor-suppressor P16/Cdkn2/Mts1 in human cancers. Nat Med 1995;1(7):686–92.
[5] Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, et al. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. Nat Genet 2012;44(1):40–6.
[6] Sproul D, Meehan RR. Genomic insights into cancer-associated aberrant CpG island hypermethylation. Brief Funct Genomics 2013;12(3):174–90.
[7] Koslowski M, Luxemburger U, Tureci O, Sahin U. Tumor-associated CpG demethylation augments hypoxia-induced effects by positive autoregulation of HIF-1alpha. Oncogene 2011;30(7):876–82.
[8] Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet 2011;20(4):670–80.
[9] Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 2010;466(7303):253–7.
[10] Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. Oncotarget 2012;3(4):462–74.
[11] Finnegan DJ. Eukaryotic transposable elements and genome evolution. Trends Genet 1989;5(4):103–7.
[12] Pace II JK, Feschotte C. The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. Genome Res 2007;17(4):422–32.
[13] Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 1998;20(2):116–7.
[14] Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 2004;431(7004):96–9.
[15] Hata K, Sakaki Y. Identification of critical CpG sites for repression of L1 transcription by DNA methylation. Gene 1997;189(2):227–34.
[16] Lin CH, Hsieh SY, Sheen IS, Lee WC, Chen TC, Shyu WC, et al. Genome-wide hypomethylation in hepatocellular carcinogenesis. Cancer Res 2001;61(10):4238–43.
[17] Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, et al. Increased methylation variation in epigenetic domains across cancer types. Nat Genet 2011;43(8):768–75.
[18] Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Bba-Rev Cancer 2007;1775(1):138–62.
[19] Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature 1997;389(6652):745–9.

[20] Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 2000;405(6785):482–5.

[21] Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science 2007;315(5815):1141–3.

[22] Mohandas T, Sparkes RS, Shapiro LJ. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. Science 1981;211 (4480):393–6. http://www.ncbi.nlm.nih.gov/pubmed/6164095. http://www.sciencemag.org/content/211/4480/393.long

[23] Bestor TH. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. Embo J 1992;11(7):2611–7.

[24] Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992;69(6):915–26.

[25] Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, et al. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development 1996;122(10):3195–205.

[26] Okano M, Xie SP, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet 1998;19(3):219–20.

[27] Margot JB, Cardoso MC, Leonhardt H. Mammalian DNA methyltransferases show different subnuclear distributions. J Cell Biochem 2001;83(3):373–9.

[28] Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99(3):247–57.

[29] Watanabe D, Suetake I, Tada T, Tajima S. Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. Mech Develop 2002;118(1–2):187–90.

[30] Pradhan S, Bacolla A, Wells RD, Roberts RJ. Recombinant human DNA (cytosine-5) methyltransferase I. Expression, purification, and comparison of de novo and maintenance methylation. J Biol Chem 1999;274(46):33002–10.

[31] Easwaran HP, Schermelleh L, Leonhardt H, Cardoso MC. Replication-independent chromatin loading of Dnmt1 during G2 and M phases. Embo Rep 2004;5(12):1181–6.

[32] Kim GD, Ni JW, Kelesoglu N, Roberts RJ, Pradhan S. Cooperation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. Embo J 2002;21(15):4183–95.

[33] Schaefer M, Pollex T, Hanna K, Tuorto F, Meusburger M, Helm M, et al. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. Gene Dev 2010;24(15):1590–5.

[34] Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 2011;333(6047):1300–3.

[35] Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science 2013;339(6118):448–52.

[36] Hur K, Cejas P, Feliu J, Moreno-Rubio J, Burgos E, Boland CR, et al. Hypomethylation of long interspersed nuclear element-1 (LINE-1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. Gut 2014;63(4):635–46.

[37] Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. Science 2009;324(5929):929–30.

[38] Zluvova J, Janousek B, Vyskot B. Immunohistochemical study of DNA methylation dynamics during plant development. J Exp Bot 2001;52(365):2265–73.

[39] Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 2005;37(8):853–62.

[40] Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, et al. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 2008;26(7):779–85.

[41] Serre D, Lee BH, Ting AH. MBD-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. Nucleic Acids Res 2010;38(2):391–9.

[42] Frommer M, Mcdonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci USA 1992;89(5):1827–31.

[43] Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci USA 1996;93(18):9821–6.

[44] Adorjan P, Distler J, Lipscher E, Model F, Muller J, Pelet C, et al. Tumour class prediction and discovery by microarray-based DNA methylation analysis. Nucleic Acids Res 2002;30(5):e21.

[45] Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics 2011;98(4):288–95.

[46] Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009;462(7271):315–22.

[47] Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res 2005;33(18):5868–77.

[48] Summerer D. Enabling technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. Genomics 2009;94(6):363–8.

[49] Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, et al. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science 2012;336(6083):934–7.

[50] Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 2012;149(6):1368–80.

[51] Rivetti C, Codeluppi S. Accurate length determination of DNA molecules visualized by atomic force microscopy: evidence for a partial B- to A-form transition on mica. Ultramicroscopy 2001;87(1–2):55–66.

[52] Oberdoerffer P, Sinclair DA. The role of nuclear architecture in genomic instability and ageing. Nat Rev Mol Cell Bio 2007;8(9):692–702.

[53] Kamakaka RT, Biggins S. Histone variants: deviants? Gene Dev 2005;19(3):295–310.

[54] Clausell J, Happel N, Hale TK, Doenecke D, Beato M. Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF. PLoS ONE 2009;4(10).

[55] Tanaka Y, Kato S, Tanaka M, Kuji N, Yoshimura Y. Structure and expression of the human oocyte-specific histone H1 gene elucidated by direct RT-nested PCR of a single oocyte. Biochem Biophys Res Commun 2003;304(2):351–7.

[56] Drabent B, Kardalinou E, Doenecke D. Structure and expression of the human gene encoding testicular H1 histone (H1t). Gene 1991;103(2):263–8.

[57] Stargell LA, Bowen J, Dadd CA, Dedon PC, Davis M, et al. Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila. Genes Dev 1993;7(12B):2641–51.

[58] Bargaje R, Alam MP, Patowary A, Sarkar M, Ali T, Gupta S, et al. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. Nucleic Acids Res 2012;40(18):8965–78.

[59] Gamble MJ, Frizzell KM, Yang C, Krishnakumar R, Kraus WL. The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. Genes Dev 2010;24(1):21–32.

[60] Chadwick BP, Willard HF. Cell cycle-dependent localization of macroH2A in chromatin of the inactive X chromosome. J Cell Biol 2002;157(7):1113–23.

[61] Nusinow DA, Hernandez-Munoz I, Fazzio TG, Shah GM, Kraus WL, Panning B. Poly(ADP-ribose) polymerase 1 is inhibited by a histone H2A variant, MacroH2A, and contributes to silencing of the inactive X chromosome. J Biol Chem 2007;282(17):12851–9.

[62] Chadwick BP, Willard HF. A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. J Cell Biol 2001;152(2):375–84.

[63] Boulard M, Gautier T, Mbele GO, Gerson W, Hamiche A, et al. The NH2 tail of the novel histone variant H2BFWT exhibits properties distinct from conventional H2B with respect to the assembly of mitotic chromosomes. Mol Cell Biol 2006;26(4):1518–26.

[64] Hake SB, Garcia BA, Duncan EM, Kauer M, Dellaire G, Shabanowitz J, et al. Expression patterns and post-translational modifications associated with mammalian histone H3 variants. J Biol Chem 2006;281(1):559–68.

[65] Loyola A, Almouzni G. Marking histone H3 variants: how, when and why? Trends Biochem Sci 2007;32(9):425–33.

[66] Hake SB, Allis CD. Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". Proc Natl Acad Sci USA 2006;103(17):6428–35.

[67] Wong LH, Ren H, Williams E, McGhie J, Ahn S, Sim M, et al. Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. Genome Res 2009;19(3):404–14.

[68] Sullivan KF, Hechenberger M, Masri K. Human Cenp-a contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol 1994;127(3):581–92.

[69] Sarma K, Reinberg D. Histone variants meet their match. Nat Rev Mol Cell Bio 2005;6(2):139–49.

[70] Moosmann A, Campsteijn C, Jansen PWTC, Nasrallah C, Raasholm M, Stunnenberg HG, et al. Histone variant innovation in a rapidly evolving chordate lineage. Bmc Evol Biol 2011;11:208.

[71] Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat Rev Genet 2008;9(11):843–54.

[72] Dominski Z, Zheng LX, Sanchez R, Marzluff WF. Stem-loop binding protein facilitates 3′-end formation by stabilizing U7 snRNP binding to histone pre-mRNA. Mol Cell Biol 1999;19(5):3561–70.

[73] Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 2008;40(7):897–903.

[74] Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, et al. Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci USA 2002;99(13):8695–700.

[75] Stewart MD, Li JW, Wong JM. Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. Mol Cell Biol 2005;25(7):2525–38.

[76] Grunstein M. Histone acetylation in chromatin structure and transcription. Nature 1997;389(6649):349–52.

[77] Dion MF, Altschuler SJ, Wu LF, Rando OJ. Genomic characterization reveals a simple histone H4 acetylation code. Proc Natl Acad Sci USA 2005;102(15):5501–6.

[78] Wang Y, Kallgren SP, Reddy BD, Kuntz K, Lopez-Maury L, Thompson J, et al. Histone H3 lysine 14 acetylation is required for activation of a DNA damage checkpoint in fission yeast. J Biol Chem 2012;287(6):4386–93.

[79] Watanabe S, Radman-Livaja M, Rando OJ, Peterson CL. A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme. Science 2013;340(6129):195–9.

[80] Chen X, Lu CG, Prado JRM, Eun SH, Fuller MT. Sequential changes at differentiation gene promoters as they become active in a stem cell lineage. Development 2011;138(12):2441–50.

[81] Cheung WL, Ajiro K, Samejima K, Kloc M, Cheung P, Mizzen CA, et al. Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. Cell 2003;113(4):507–17.

[82] Qian JB, Lesage B, Beullens M, Van Eynde A, Bollen M. PP1/repoman dephosphorylates mitotic histone H3 at T3 and regulates chromosomal aurora B targeting. Curr Biol 2011;21(9):766–73.

[83] Jenuwein T, Allis CD. Translating the histone code. Science 2001;293(5532):1074–80.

[84] Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell 2007;129(4):823–37.

[85] Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006;125(2):315–26.

[86] Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 2005;123(4):581–92.

[87] Collas P. The current state of chromatin immunoprecipitation. Mol Biotechnol 2010;45(1):87–100.

[88] Rhee HS, Pugh BF. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell 2011;147(6):1408–19.

[89] van Steensel B, Henikoff S. Identification of *in vivo* DNA targets of chromatin proteins using tethered dam methyltransferase. Nat Biotechnol 2000;18(4):424–8.

[90] Vogel MJ, Peric-Hupkes D, van Steensel B. Detection of *in vivo* protein-DNA interactions using DamID in mammalian cells. Nat Protoc 2007;2(6):1467–78.

[91] Zhou V. Methods for global characterization of chromatin regulators in human cells. Harvard University Press; 2012.

[92] Kelly TK, Liu Y, Lay FD, Liang G, Berman BP, Jones PA. Genome-wide mapping of nucleosome positioning and DNA methylation in individual DNA molecules. Genome Res 2012;22(12):2497–506.

[93] Matarazzo MR, Lembo F, Angrisano T, Ballestar E, Ferraro M, Pero R, et al. *In vivo* analysis of DNA methylation patterns recognized by specific proteins: coupling CHIP and bisulfite analysis. Biotechniques 2004;37(4):666–8. 70, 72–3.

[94] Cassidy SB, Dykens E, Williams CA. Prader-Willi and Angelman syndromes: sister imprinted disorders. Am J Med Genet 2000;97(2):136–46.

[95] Skryabin BV, Gubar LV, Seeger B, Pfeiffer J, Handel S, Robeck T, et al. Deletion of the MBII-85 snoRNA gene cluster in mice results in postnatal growth retardation. PLoS Genet 2007;3(12):e235.

[96] de Smith AJ, Purmann C, Walters RG, Ellis RJ, Holder SE, Van Haelst MM, et al. A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism. Hum Mol Genet 2009;18(17):3257–65.

[97] Runte M, Varon R, Horn D, Horsthemke B, Buiting K. Exclusion of the C/D box snoRNA gene cluster HBII-52 from a major role in Prader-Willi syndrome. Hum Genet 2005;116(3):228–30.

[98] Moulton T, Crenshaw T, Hao Y, Moosikasuwan J, Lin N, Dembitzer F, et al. Epigenetic Lesions at the H19 Locus in Wilms-Tumor Patients. Nat Genet 1994;7(3):440–7.

[99] Weksberg R, Shen DR, Fei YL, Song QL, Squire J. Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. Nat Genet 1993;5(2):143–50.

[100] Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi CF, Wolffe A, et al. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. Curr Biol 2000;10(14):853–6.

[101] Robertson KD. DNA methylation and human disease. Nat Rev Genet 2005;6(8):597–610.

[102] Cui H, Niemitz EL, Ravenel JD, Onyango P, Brandenburg SA, Lobanenkov VV, et al. Loss of imprinting of insulin-like growth factor-II in Wilms' tumor commonly involves altered methylation but not mutations of CTCF or its binding site. Cancer Res 2001;61(13):4947–50.

[103] Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 1983;301(5895):89–92.

[104] Gamasosa MA, Slagel VA, Trewyn RW, Oxenhandler R, Kuo KC, Gehrke CW, et al. The 5-methylcytosine content of dna from human-tumors. Nucleic Acids Res 1983;11(19):6883–94.

[105] Jones PA. DNA methylation and cancer. Oncogene 2002;21(35):5358–60.

[106] Suter CM, Martin DI, Ward RL. Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue. Int J Colorectal Dis 2004;19(2):95–101.

[107] Gonzalgo ML, Jones PA. Mutagenic and epigenetic effects of DNA methylation. Mutat Res-Rev Mutat 1997;386(2):107–18.

[108] Christman JK. 5-Azacytidine and 5-aza-2′-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. Oncogene 2002;21(35):5483–95.

[109] Thakur S, Feng XL, Shi ZQ, Ganapathy A, Mishra MK, Atadja P, et al. ING1 and 5-azacytidine act synergistically to block breast cancer cell growth. PLoS One 2012;7(8).

[110] Walker FO. Huntington's disease. Lancet 2007;369(9557):218–28.

[111] Steffan JS, Bodai L, Pallos J, Poelman M, McCampbell A, Apostol BL, et al. Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. Nature 2001;413(6857):739–43.

[112] Sadri-Vakili G, Cha JHJ. Mechanisms of disease: histone modifications in Huntington's disease. Nat Clin Pract Neuro 2006;2(6):330–8.

[113] Fischer A, Sananbenesi F, Wang XY, Dobbin M, Tsai LH. Recovery of learning and memory is associated with chromatin remodelling. Nature 2007;447(7141):178-U2.

[114] Richon VM, Sandhoff TW, Rifkind RA, Marks PA. Histone deacetylase inhibitor selectively induces p21(WAF1) expression and gene-associated histone acetylation. Proc Natl Acad Sci USA 2000;97(18):10014–9.

[115] Dang WW, Steffen KK, Perry R, Dorsey JA, Johnson FB, Shilatifard A, et al. Histone H4 lysine 16 acetylation regulates cellular lifespan. Nature 2009;459(7248):802-U2.

[116] Greer EL, Maures TJ, Hauswirth AG, Green EM, Leeman DS, Maro GS, et al. Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in C. elegans. Nature 2010;466(7304):383-U137.

[117] Maures TJ, Greer EL, Hauswirth AG, Brunet A. The H3K27 demethylase UTX-1 regulates C. elegans lifespan in a germline-independent, insulin-dependent manner. Aging Cell 2011;10(6):980–90.

[118] Chen F, Kan H, Castranova V. Methylation of lysine 9 of histone H3: role of heterochromatin modulation and tumorigenesis. Handbook of Epigenetics: The New Molecular and Medical Genetics; 2011. 149–57.

[119] Pedersen MT, Helin K. Histone demethylases in development and disease. Trends Cell Biol 2010;20(11):662–71.

[120] Morin RD, Johnson NA, Severson TM, Mungall AJ, An JH, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet 2010;42(2):181-U24.

[121] Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 2002;419(6907):624–9.

[122] Crea F, Fornaro L, Bocci G, Sun L, Farrar WL, Falcone A, et al. EZH2 inhibition: targeting the crossroad of tumor invasion and angiogenesis. Cancer Metast Rev 2012;31(3–4):753–61.

[123] Sarraf SA, Stancheva I. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA—replication and chromatin assembly. Mol Cell 2004;15(4):595–605.

[124] Tamaru H, Selker EU. A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature 2001;414(6861):277–83.

[125] Jackson JP, Lindroth AM, Cao XF, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature 2002;416(6880):556–60.

This page is intentionally left blank

Part III

Experimental and Discovery Platforms

This page is intentionally left blank

Chapter 5

Animal Models in Pharmacogenomics

Erin Olson and Delyth Graham

BHF Glasgow Cardiovascular Research Centre, University of Glasgow, UK

Chapter Outline

5.1 Overview 73
5.2 Rodent Models 74
  5.2.1 Mouse Models 74
  5.2.2 Rat Models 74
  5.2.3 Outbred Rodent Stocks 74
  5.2.4 Inbred Strains 75
  5.2.5 Heterogeneous Stocks 75
  5.2.6 Collaborative Cross 75
  5.2.7 Recombinant Inbred Strains 76
  5.2.8 Congenic Strains 76
  5.2.9 Consomic Strains 77
  5.2.10 Transgenic Rodent Models 77
5.3 Nonrodent Models of Human Disease 79
5.4 Genetic/Genomic Resources 79
5.5 Ethical Issues of Animal Experimentation 80
  5.5.1 Replacement 80
  5.5.2 Refinement 80
  5.5.3 Reduction 80
5.6 Caution and Limitations 80
5.7 Applications of *in vivo* Models in Pharmacogenomic Research 81
  5.7.1 Quantitative Trait Mapping 81
  5.7.2 Identification of eQTLs 81
5.8 Rodent Models in Pharmacogenomics 82
  5.8.1 Bipolar Disorders 82
  5.8.2 Anxiety Disorders 82
  5.8.3 Pharmacogenomics Using Xenograft Mouse Models of Acute Leukemia 82
  5.8.4 Pharmacogenomics of Pain and Analgesia 83
  5.8.5 Pharmacogenomics and PAH 84
  5.8.6 Pharmacodynamic/Pharmacogenomic Modeling of Insulin Resistance 84
5.9 Conclusion 85
References 85

5.1 OVERVIEW

Pharmacogenomics is the application of genome-wide approaches for determining genetic influences on drug response and for the development of novel therapies. By investigating the gene expression profile of an individual disorder, specific therapy can be targeted to that disorder, leading to greater efficacy and reduced toxicity [1]. Recent advances in high-throughput methods of generating biological datasets and in the computational power and bioinformatics needed to manipulate these vast datasets have completely transformed the manner in which this research is conducted. Where human studies are ethically or practically impossible, animal models can play an important role in pharmacogenomic investigations.

This chapter discusses the approaches available to exploit high-throughput technology and the genetic diversity of rat and mouse species in mechanistic pharmacogenomics research. It also provides evidence to show that the use of appropriate genetically defined rodent models not only aids identification of genetic determinants of drug susceptibility but also helps to establish underlying mechanisms.

Human pharmacogenomic studies can be particularly problematic in terms of tissue collection and intervention. Generally, it is not possible to obtain tissues from a wide variety of anatomical sites or during different developmental stages in humans. Additionally, it is ethically unacceptable to expose humans to toxic substances or carry out genetic engineering. While cell-based assays can provide a useful

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00005-0
© 2014 Elsevier Inc. All rights reserved.

alternative for generating pharmacogenomic datasets and have been used extensively in preclinical drug development, *in vitro* studies examine pharmacology in isolation from adaptive and regulatory networks and signals that exist in the whole organism [2]. Thus a great importance is attributed to experimental animal models. Currently the most utilized models are the mouse and the rat.

5.2 RODENT MODELS

Genetic studies using rodents can circumvent many of the limitations commonly encountered with human studies and *in vitro* approaches in pharmacogenomics. Inbred rodent models allow studies to be conducted on a homogeneous genetic background, thereby reducing the noise generated from human heterogeneity. Moreover, in rodent studies gene–environment interactions are minimized, well-defined, and well-controlled when compared to human studies [3]. The traditional role of rat and mouse models has been further supported by the sequencing of their genomes, where rodent models have been shown to have greater evolutionary relatedness to humans than expected [4].

Rats are commonly used in drug-safety evaluation experiments, and genetically diversified inbred strains are being developed and used for population-based toxicity testing [5]. However, the most widely used laboratory model for pharmacogenomic studies is the mouse, which, until very recently, offered an unparalleled wealth of genetic knowledge and resources, including a high-density SNP database containing more than 8 million polymorphic loci in hundreds of inbred strains [6–8].

5.2.1 Mouse Models

There are many advantages to using mice as research models: specifically, their genetic characterization, the large number of strains available, and the large list of cataloged mutant genes that provide models suited to a number of different areas of research. Mice are relatively inexpensive to breed and maintain compared to other species. They are also easy to handle and care for. Their high reproductive performance, large litter size, and short gestation means that many generations can be produced in a relatively short period of time (with the potential for 1 million descendants after 425 days).

The disadvantages of mice as research animals include their small size, which can limit the procedures that may be performed, and the sample volume size that can be obtained from an individual animal. In order to overcome volume size limitations, samples from several animals may be pooled for research analysis and for improved statistical power.

The use of the mouse as a research animal has resulted in many scientific advances. Much of our early understanding of the immune system resulted from mouse studies [9]. The use of the mouse continues to be an important part of various fields of research, including embryology, aging, cancer, pharmacological and toxicological testing, and infectious disease. Transgenic and knockout mice have become important tools for investigating the relationship of genetic makeup to disease states as well as for elucidating pathways of normal mammalian development [10].

5.2.2 Rat Models

The laboratory rat was the first mammalian species domesticated for scientific research, and it has been a well-studied animal model across the biomedical sciences. Inbred rat strains were first established by Helen Dean King nearly a century ago [11], and they are used for research into a broad array of human conditions and diseases, such as addiction, aging, autoimmunity, cardiovascular disease, hypertension, metabolic disorders, and cancer. Because of its well-studied physiology and the relative ease of experimental intervention, the rat is also widely used in the pharmaceutical industry to assess the efficacy and toxicity of drugs.

While the rat has been primarily known as an excellent "physiological" model, it has lagged behind the mouse as a genetic model in the endeavor to identify genetic contributions to disease. In recent years, however, there has been a steady increase in the use of the rat for genomic and genetic studies [12,13]. The use of the rat for genetic/genomic research was transformed by the sequencing of the rat genome in 2004 [4], which provided an exceptional opportunity for identifying genes and pathways underlying disease phenotypes. In evolutionary terms, the rat is approximately 5 million years closer to the human than the mouse, thus allowing greater translational opportunities [14].

There are three main genetic types of rodent used in biomedical research. These are (1) animals that come from a genetically heterogeneous (nonisogenic) group, such as an outbred stock, an F_2 hybrid, or a genetically heterogeneous colony, (2) animals that come from a genetically uniform (isogenic) group, including inbred strains and F_1 hybrids; and (3) animals expressing a major gene mutation.

5.2.3 Outbred Rodent Stocks

Outbred rodent stocks (e.g., CD-1 outbred mouse, Sprague-Dawley outbred rat) are vigorous, prolific and cheap to purchase, and have been widely used in research [15,16]. They are usually maintained as large breeding colonies within which there is inter-individual genetic variation. They are maintained by random (or haphazard) mating systems. Each animal will be genetically different, but the extent of genetic variation depends on the parental strain origins of the colony. For example, when a new colony is established or during caesarean rederivation for the elimination of disease, genetic variation will be reduced whereas an outcross to a different stock will increase variability.

As research models, outbred stocks have a number of disadvantages. For one, their characteristics can change rapidly due to selection, inbreeding, and random genetic drift. Also, outbred animals are usually significantly heavier than inbred ones as a result of many generations of selection for fast growth rate and large litter sizes. They are additionally "genetically undefined"—that is, the genotype of an individual is unknown unless it is specifically assessed. Stock names such as "Sprague-Dawley," "Wistar," or "Swiss" have little genetic meaning, and there are no genetic markers to define them. Outbred stock from different breeders will be genetically different (due to different husbandry regimes), which means that historical data collected on such stocks may be unreliable, making quality control difficult.

The use of outbred stock is often justified on the grounds that "humans are outbred," which should make it easier to "extrapolate" to humans [16]. However, except for certain specialized uses (e.g., heterogenous stock for sub-centimorgan fine mapping of quantitative trait loci or the Collaborative Cross, a resource for mammalian systems genetics), outbred stocks are not the most suitable model for genetic research. Their use leads to a larger number of false negative results because phenotype variability inevitably leads to lower-powered experiments. The almost universal use of outbred stocks in toxicity testing has arisen by historical accident and has not been scientifically justified.

5.2.4 Inbred Strains

Inbred strains are produced by greater than 20 generations of brother × sister mating, with all individuals tracing back to a single pair in the twentieth or subsequent generations [17]. They are genetically stable and cannot be changed by selective breeding. However, sublines have arisen in most of the commonly used strains as a result of residual heterozygosity (i.e., the sublines were separated before the strain was fully inbred) and new mutations, which can result in genetic drift within geographically isolated sublines of the same inbred strain. There are more than 400 inbred strains of mice and 150 inbred strains of rats, and geneticists have recognized their value for many years.

The key characteristics of inbred strains are as follows:

- Isogenicity: all individuals within a strain are genetically identical, so the same genotype can be repeatedly reproduced.
- Homozygosity: animals are homozygous at virtually all genetic loci. This leads to immortality of the genotype because offspring are genetically identical to their parents.
- Phenotypic consistency: uniformity of the phenotype means greater statistical power and the requirement for fewer animals per experimental group.
- Long-term stability: once fully inbred, these strains are not changed by selective breeding. New mutations will lead to gradual genetic drift, so it is important for investigators to specify the substrains they use.
- Identifiability: each inbred strain has a unique set of genetic markers that can be used for genetic quality control, allowing regular marker screening to ensure maintenance of homozygosity within the colony.
- Individuality. each strain is unique and has particular phenotypic characteristics that are important for addressing specific research questions. Studies using particular inbred strains of mice (e.g., C57BL/6) or rats (e.g., Dahl SS) are well-accepted research strategies provided it is clearly understood that the results may apply only to that strain and may not be relevant to other strains.

Inbred strains represent fixed, renewable genotypes that are ideally suited to systems biology approaches and to identification of whole-genome response signatures.

5.2.5 Heterogeneous Stocks

Both rat and mouse heterogeneous stocks (HS) have been generated as alternative models for fine mapping of small-effect quantitative trait loci (QTLs) [18–20]. These genetically heterogeneous rodent models are outbred animals of known ancestry, descended from eight inbred progenitor strains, and outcrossed to maintain diversity. The stock is maintained for more than 50 generations so that each chromosome is a fine-grained mosaic of the progenitor strains. The average distance between recombinants is small (less than 2 centimorgans) so that the HS are a powerful tool for high-resolution mapping of multiple QTLs across the genome.

5.2.6 Collaborative Cross

The collaborative cross (CC) is a large panel of mouse recombinant-inbred (RI) strains derived from a genetically diverse set of founder strains and designed specifically for complex trait analysis [21–24]. The CC is characterized by genome-wide genetic variation, randomization, and a reproducible stable population that is sufficiently large to support statistical analysis. The eight founder strains (A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/HlLtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ) were bred through an eight-way "funnel" breeding design established to randomly mix the variation in the founder strains before inbreeding by brother × sister mating (Figure 5.1). The founder strains were chosen in order to maximize genetic diversity and utility for studying traits of widespread interest. The eight collaborative cross-founders capture ~90% of allele diversity across all 1 Mb intervals spanning the mouse genome. This is compared to the most commonly used RI mouse strains

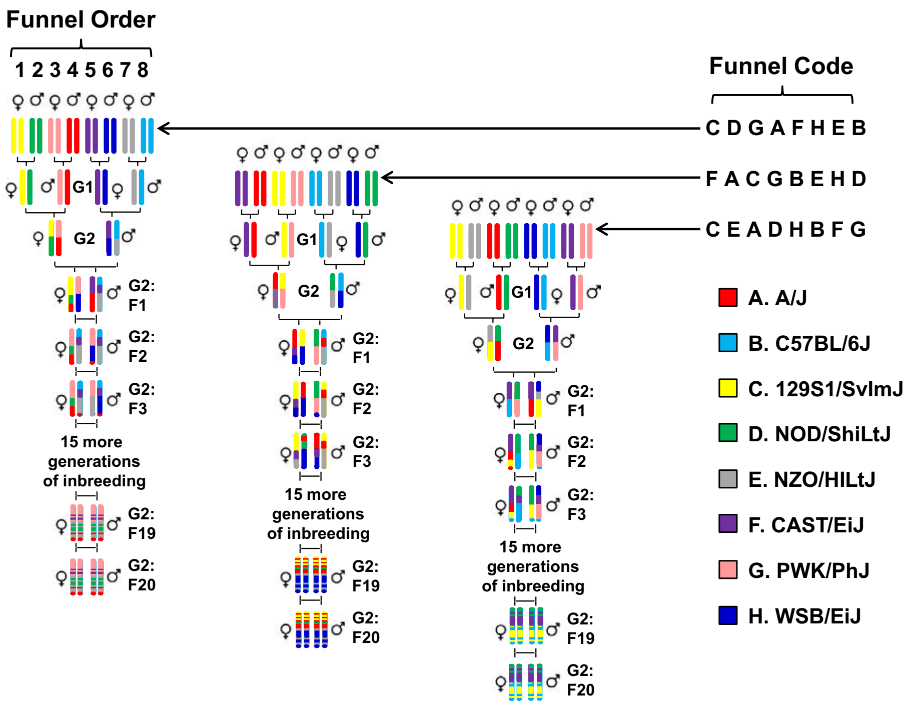


FIGURE 5.1 **Breeding scheme of CC lines.** The figure shows the breeding scheme for three independent CC lines. Each line has a funnel section followed by an inbreeding section. The eight founder strains are arranged in different positions (1–8) in each line, and this order determines the funnel code on the basis of a single letter code for each line. Founder order is randomized and not repeated across lines. Each mouse is represented by a pair of homologous autosomes and a symbol denoting its sex. *Source: Adapted from the Collaborative Cross Consortium* [24].

(AXB/BXA and BXD), which each capture only approximately 13% of known variation. The CC has a population structure that randomizes existing genetic variation, providing unparalleled power to determine causality and investigate detailed biological networks underlying disease and drug response [24].

5.2.7 Recombinant Inbred Strains

Rat and mouse recombinant-inbred strains (RIS) provide powerful tools for genetic mapping [25,26]. An RIS is formed by crossing two inbred strains followed by repeated (>20 generations) brother × sister mating to create a new inbred line whose genome is a mosaic of the parental genomes (Figure 5.2). Since each RIS is an inbred strain, it can be generated perpetually. A panel of RIS has a number of advantages for genetic mapping. These include the following: genotyping of each strain is required only once; multiple individuals from each strain may be phenotyped to reduce individual, environmental, and measurement variability; multiple phenotypes can be assessed on the same genetic background. Also, as recombination events are more dense in RIS than those occurring in any one meiosis, greater mapping resolution can be achieved.

5.2.8 Congenic Strains

Congenic rats and mice are generated to obtain strains that differ by a specific chromosomal region, which often contains only a few specified genes, or ideally by a single gene (Figure 5.2) [27]. Two inbred lines are crossed to make an F_1 generation, which is then backcrossed to one of the parental inbred strains. Backcrossing takes place for at least 10 generations in order to remove contaminating background heterozygosity while maintaining heterozygosity within the chromosomal region of interest. After at least 10 backcross generations, brother × sister mating is carried out to achieve homozygosity within the congenic interval. A speed congenic strategy can be used to significantly reduce the time taken to produce a congenic strain (from 4 years to 18 months) [28]. This is achieved by genotyping each backcross generation with 60–100 microsatellite markers evenly spaced throughout the background genome and then choosing animals with the lowest heterozygosity from which to breed and produce the next backcross generation.

Congenic strains are powerful tools for confirming the existence of QTLs identified by genome-wide scans, and production of congenic substrains allows dissection of the region of interest and aids functional studies of candidate causal genes.

5.2.9 Consomic Strains

Consomic strains are produced by repeated backcrossing of a whole chromosome, such as X or Y, onto an inbred strain. They are generated using a strategy similar to that used for congenic strain production (Figure 5.2) [27,29].

5.2.10 Transgenic Rodent Models

Mice and (more recently) rats can be genetically modified using a number of different approaches. These approaches include gene knockout, gene overexpression, conditional/inducible gene overexpression/inactivation, and tissue and/or time specific expression or knockdown [30].

Transgenic models are important tools for studying the function or regulation of a specific gene of interest and for modeling human diseases caused by dominantly acting mutant proteins (e.g., Alzheimer's disease) [31].

There are two main methods for generation of transgenic rodent models [10,31]. The first is pronuclear microinjection, where foreign DNA is introduced directly into the egg, just after fertilization. The DNA is microinjected into the large sperm-derived pronucleus, where it integrates as multiple copies at a random position in the genome, often after one or two cell divisions have occurred. If the transgenic cells contribute to the germ line, some transgenic eggs or sperm will be produced and the next generation will be fully transgenic (Figure 5.3).

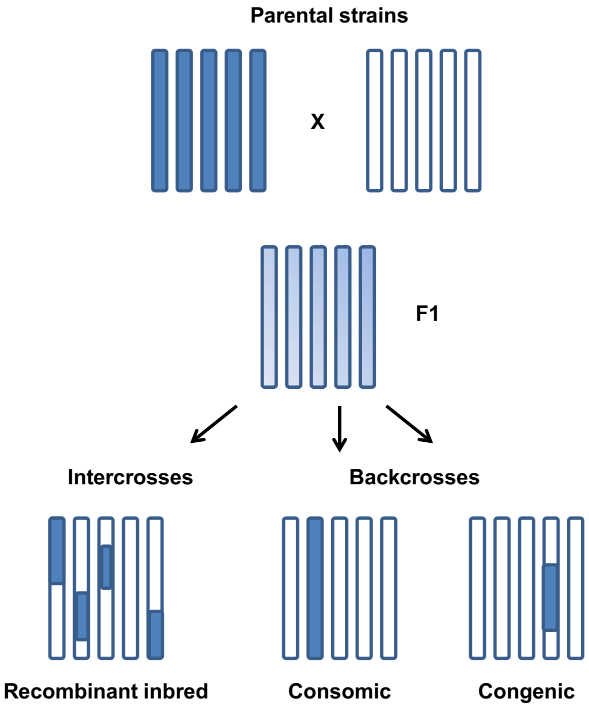


FIGURE 5.2 **Breeding strategies.** These strategies are used to generate recombinant inbred, consomic, and congenic strains.

The second method is the introduction of DNA into embryonic stem cells (ES cells). These are derived from very-early-stage embryos that can differentiate into all types of cell when introduced into another embryo. Similarly to pronuclear microinjection, DNA introduced into ES cells may integrate randomly within the genome. However, if the introduced DNA is similar in sequence to part of the host genome, it may undergo "homologous recombination" and integrate as a single copy at a specific site. ES cells colonize the host embryo and often contribute to the germ line, resulting in the production of some sperm carrying the foreign DNA. When these transgenic sperm fertilize a normal egg, transgenic offspring are produced with the same newly introduced DNA in every cell.

5.2.10.1 *Cre-LoxP Technology*

Because of early embryonic lethality, it is not always possible to generate viable transgenic mice using conventional knockout techniques. To overcome this problem, the Cre-Lox recombinase system was developed [32]. Cre-Lox recombination is a sophisticated site-specific recombinase technology that allows DNA modification to be targeted to a specific cell type or to be triggered by a specific external stimulus. Cre is a 38 kDa recombinase protein from bacteriophage P1 that catalyzes recombination between LoxP sites. LoxP sites are specific sequences of 34 base pairs (bp) consisting of an 8-bp core sequence, where recombination takes place, and two flanking 13-bp inverted repeats [31]. Since Cre gene and LoxP sites are not native to the mouse genome, they have to be introduced by transgenic (ES cell) technology [32]. Typically Cre and LoxP strains are developed separately and crossed to produce a Cre-Lox strain.

There are four main categories of Cre and LoxP strains of mice: (1) Cre-expressing strains, which contain a transgene that expresses Cre under the control of a widespread (general) or tissue-specific (conditional) promoter. They are used to produce general or conditional knockouts, respectively; (2) inducible Cre strains containing a transgene that expresses a modified form of Cre recombinase that is nonfunctional until an inducing agent (e.g., doxycycline, tetracycline, RU486, or tamoxifen) is administered at a specific time; (3) LoxP-flanked (floxed) strains, which contain LoxP sites flanking (on each side of) a critical portion of a target gene or genomic region of interest; (4) Cre reporter strains, which contain LoxP sites in combination with fluorescent or lacZ marker proteins used to determine (visualize) successful recombination and/or alterations in gene expression.

5.2.10.2 *Recent Transgenic Advances*

Zinc-finger nuclease (ZFN) technology directly allows targeted mutagenesis at the 1-cell-stage embryo; therefore, embryonic stem (ES) cells are not required. This process makes model generation much faster and avoids background

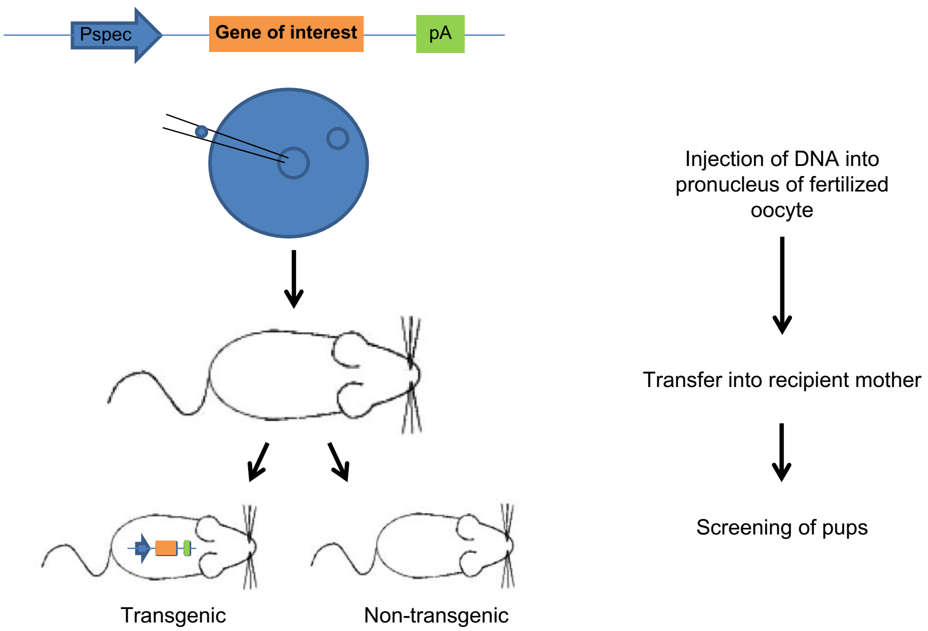

Pspec
Gene of interest
pA
Injection of DNA into pronucleus of fertilized oocyte
Transfer into recipient mother
Screening of pups
Transgenic
Non-transgenic

FIGURE 5.3 **Producing transgenic mice.** The schematic shows a transgenic construct containing a specific promoter (Pspec), the gene of interest, and a polyadenylation signal (pA). This construct is injected into fertilized oocytes to be subsequently transferred into pseudo-pregnant recipient mothers. Pups are then screened for integration of the transgene. *(Modified from Bockamp E, Maringer M, Spangenberg C, Fees S, Fraser S, Eshkind L, Oesch F, Zabel B. Of mice and models: improved animal models for biomedical research. Physiol Genomics. 2002 Dec 3;11(3):115–32.)*

strain limitations. Unlike ES cell transgenesis, ZFN technology is directly applicable to the rat and has been used in recent years to successfully create numerous transgenic rat models on a variety of genetic backgrounds [33,34].

Zinc-finger proteins are well-characterized DNA binding proteins that, for transgenic purposes, have been coupled with a Fok1 endonuclease and designed to recognize and cleave specific DNA sequences. ZFNs are utilized to generate sequence-specific double-strand breaks that are repaired by error-prone nonhomologous end joining (NHEJ) or high-fidelity homologous recombination (HR) [35,36]. Specific DNA cleavage is achieved by varying the zinc-finger domains, as each zinc finger interacts with a particular triplet of DNA base pairs. Combining different zinc fingers permits specific binding to 9–12 bp motifs. Double-stranded breaks occur when two ZFNs bind to the target DNA, bringing their nuclease domains together. One early example of successful ZFN-mediated genetic disruption in the rat includes targeting to endogenous immunoglobulin M (IgM) and Rab38 genes as well as to an integrated GFP reporter gene [33]. ZFN technology is also successfully utilized in *Drosophila* and zebrafish [36].

More recently, transcription activator–like effector nucleases (TALENS) have been used for efficient targeted genome editing. TALENS are similar to ZFNs, comprising a nonspecific Fok1 nuclease domain fused to a customizable DNA-binding domain. They are easily and rapidly designed and have been shown to have high success rates, which makes them suitable for use by nonspecialist researchers [37,38].

Other strategies for producing knockout animals include chemical mutagenesis and the use of mobile DNA elements (transposons and retrotransposons). Chemical mutagenesis using N-ethyl-N-nitrosourea (ENU) is a method of random mutagenesis used in mice and rats [39–41]. ENU is an alkylating agent that transfers its ethyl group to nitrogen or oxygen radicals in DNA, resulting in base mispairing that causes base pair substitution. ENU mutagenesis creates one base pair change in any given gene in every 200–700 gametes. ENU is injected into male mice that are then bred with wild-type females to produce mutant offspring. The offspring are phenotyped, and, once recognized, the causative mutation is mapped by positional cloning [40,42].

Another strategy for targeted gene knockdown is by RNA interference (RNAi) [43]. RNAi is a highly conserved, naturally occurring mechanism of gene suppression found in plant, yeast, and mammalian cells. It has evolved as a valuable biological tool for gene identification and for studying post-transcriptional gene function. RNAi is mediated by naturally occurring endogenous or synthetic exogenous short interfering RNAs (siRNAs), which typically consist of 21–24-nucleotide duplexes. The antisense siRNA strand is incorporated into a RNA-induced silencing complex (RISC), which then targets complementary mRNA sequences in the cytoplasm. mRNA is then cleaved by RNase at sites not bound by siRNA, thus preventing protein translation [44]. siRNA offers many unique features over traditional methods of inhibition (i.e., chemical inhibitors), including sequence specificity, high efficacy, and the ability to target a large number of genes; however, one drawback is its limited delivery to target cells *in vivo*. Unprotected or naked siRNA is able to penetrate cell membranes when injected locally or systemically, but it is rapidly broken down by plasma nucleases [45]. One method shown to improve systemic delivery of naked siRNA to target organs is hydrodynamic tail vein injection, which involves rapid delivery of a large volume of siRNA plus transfection reagent by tail vein or femoral vein. Alternative *in vivo* delivery methods include viral and bacterial vectors, nanoparticle technology, and liposomes, which have all been used with varying success. Differential uptake by various organs depends on the delivery technique, and this should be taken into account when deciding on the most appropriate method of delivery for *in vivo* siRNA.

5.3 NONRODENT MODELS OF HUMAN DISEASE

The zebrafish is becoming an increasingly popular model organism in pharmacogenomics and network analysis. It possesses many features that make it a highly useful model system. Adults are relatively small, inexpensive, and easily housed and maintained in large quantities. They are easily mated under laboratory-controlled conditions, with each pair producing several hundred eggs per mating. The embryos are transparent, allowing easy visualization of fundamental developmental processes. They develop rapidly, and by 5 days post-fertilization larvae have developed the majority of their internal organs, including the liver, pancreas, and a complex vascular network. This relatively short generation time makes genetic screens easier, faster, and cheaper than they are with mice. Thus the zebrafish is currently one of the main model systems for drug discovery and validation. Novel methods to validate potential drug targets in zebrafish include morpholino antisense knockdown technology and target-selected mutagenesis [46].

A variety of human diseases have been modeled on zebrafish. For example, Tanaka et al. [46] developed a zebrafish cardiac failure model. Antisense molecules such as morpholino phosphorodiamidate oligonucleotides (MOs) and peptide nucleic acids (PNAs) have been demonstrated to effectively knock down specific gene function in zebrafish. Translation-blocking MOs interfere with the function of spliceosomes, leading to exon skipping or the use of cryptic splice sites, which results in aberrantly spliced mRNA. The development of *in vivo* assays that are related to human heart failure may lead to the discovery of novel potential drug targets.

In this section, a wide variety of rodent and nonrodent models of human disease have been described, together with a range of applicable gene modification techniques. The choice of gene modification approach will depend, in part, on the gene and organ system of interest and on the relevance of the animal model's biochemistry and physiology to that of the human disease under investigation.

5.4 GENETIC/GENOMIC RESOURCES

There are extensive, publicly available databases on mouse, rat, and zebrafish genetics. Within the research community, distribution and sharing of genetic/genomic datasets is actively encouraged.

A list of the most widely used databases is as follows:

- The Mouse Phenome Database has data on a wide range of strain characteristics, searchable by subject area (e.g., behavior, blood, bone, development), strain, intervention, study design, and so forth (http://www.phenome.jax.org).
- The Mouse Genome Informatics database has data on genes, phenotypes, disease models, gene expression, gene function, pathways, recombinases, strains and SNPs, tumors, and orthology (http://www.informatics.jax.org).
- The International Mouse Strain Resources (IMSR) database is a searchable database of mouse stocks and strains available worldwide (http://www.findmice.org).
- The JAX mice database provides extensive information on mouse genetics, specifically relating to mice maintained by the Jackson Laboratory, Bar Harbor, Maine (http://www. jaxmice.jax.org).
- The Rat Genome Database (RGD) provides a comprehensive database on rat genetics. (http://rgd.mcw.edu, RGD). RGD curates and integrates rat genetic and genomic data and provides free access to support research using the rat as a genetic model for the study of human disease. RGD works in collaboration with the Mouse Genome Database (MGI), NCBI, UCSC, EBI, RGSC, Baylor College of Medicine, SWISS-PROT, BIND, RatMap, and the Phys-Gen PGA [47].
- The NBRP-Rat Phenome Database provides functional characterization (phenome database) for deposited rat strains at the National Bio-Resource Project, Japan. (http://www.anim.med.kyoto-u.ac.jp/nbr/phenome.aspx).
- The Zebrafish Neurophenome Database (ZNP) (http://www.tulane.edu/~znpindex/).
- The European Zebrafish Resource Centre is a public database of available mutant stocks that can be immediately requested from a link through the model organism database, ZFIN (http://zf-health.org/data/databases.html).
- All data from initial phenotyping of knockout mutants is made publicly available through the website of the Zebrafish Mutation Project. Transcript counting results

are publicly available through a custom interface based on ArrayExpress. Information on enhancer elements, synteny blocks, and epigenetic marks can be obtained for any region of the human genome through Ancora, a custom genome browser.
- All information from gene expression mapping of the brain is released through a custom-designed online neuroanatomy atlas, ZebrafishBrain.org (developed in the ZF-MODELS project) and through other databases.

5.5 ETHICAL ISSUES OF ANIMAL EXPERIMENTATION

Since 1986, the European Union has had in place specific legislation covering the use of animals for scientific purposes. On September 22, 2010, the EU adopted Directive 2010/63/EU, which updated and replaced the 1986 directive, 86/609/EEC, on the protection of animals used for scientific purposes. The aim of the new directive was to strengthen legislation and improve the welfare of animals needed for scientific research, as well as to firmly anchor the principle of the Three Rs (Replace, Reduce, and Refine the use of animals) in EU legislation. Directive 2010/63/EU took full effect on January 1, 2013.

In the United Kingdom, an ethical framework for the use of animals in research has been written into law by implementation of the Animals (Scientific Procedures) Act 1986. It exceeded the requirements of the European Union's Directive 86/609/EEC on the protection of animals used for experimental and other scientific purposes. The U.K. act requires that proposals for research involving animals be fully assessed in terms of any harm to the animals. This involves detailed examination of the procedures and experiments and the numbers and types of animal to be used. The proposed experimental procedures are then weighed against the potential benefits of the project. Before the introduction of Directive 2010/63/EU, this cost–benefit analysis was almost unique to U.K. animal research legislation; only German law had a similar requirement. However, the adoption of Directive 2010/63/EU has harmonized legislation across Europe [48].

Before undertaking any research involving animal models in the United Kingdom, it is an absolute requirement that the three R's be considered to fully justify the plan of work and to obtain approval of the project proposal by both local ethics committees and the Home Office Inspector [49,50] ([49,50]). (For further information, the websites http://www.nc3rs.org.uk and http://www.homeoffice.gov.uk/comrace/animals/further info.html may be consulted) The three Rs are as follows.

5.5.1 Replacement

Wherever possible live animals should be replaced by non-sentient or less sentient alternatives such as cell cultures, invertebrates, or mathematical models. There are many examples showing that this can be successful:

- Rabbits and mice were previously used to assay batches of insulin, but now this can be done chemically.
- Monoclonal antibodies were grown as ascites tumors in mice, but this is now done *in vitro*.
- In toxicity testing, there are a number of tests, such as the Ames test for mutagenesis, that, at least partially, replace studies in live animals.

5.5.2 Refinement

Pain, distress or lasting harm should be minimized.

- Anaesthesia and analgesia should be used in surgery.
- Humane end points should be used when death is the expected outcome.
- Tumors should not be allowed to grow to an excessive size.
- Animals should be protected from disease.
- An enriched environment with space for natural behavior should be provided.
- Food and water should be withdrawn only for strictly limited periods.
- Social animals, such as mice and rats, should be housed with other animals (a legal requirement under Directive 2010/63/EU).

5.5.3 Reduction

Use the minimum number of animals consistent with achieving the objectives of the study:

- Develop a research strategy with clearly defined objectives that can be achieved with available resources.
- Choose an animal model capable of answering the question being investigated.
- Undertake experiments that use neither too many animals, so that some are wasted, nor too few, so that important effects are missed.
- Carry out the correct statistical analysis of the results, including summary statistics such as means and standard deviations, as well as indicators of uncertainty, such as significance levels and confidence intervals.

5.6 CAUTION AND LIMITATIONS

Animal models yield important insights into the genetic basis of human diseases, and they provide a test bed for pharmacological treatments. Nonetheless, investigators must carefully consider their choice of model: no single model perfectly recreates the human disease, and there are related considerations of cost, infrastructure, and specialized personnel. Taking these considerations into account, experimenters need to select models that best reproduce the aspect of disease being investigated. In particular, when moving from bench to bedside it is essential to test procedures in relevant models that yield highly reproducible results. These limitations notwithstanding, the range of animal models available today ensures that it will always be possible to devise an appropriate strategy where animal models remain the best tools for advancing understanding of the mechanisms of human disease and identification of appropriate therapeutics.

As pharmacogenomics using animal models becomes increasingly important in preclinical drug development, improvements to all facets of it will be made. Technology improvements will continue to increase the sensitivity, speed, and accuracy with which differentially expressed genes are identified. Genome sequencing will incrementally enhance the quality of annotated DNA and protein databases and will facilitate functional characterization of differentially expressed genes [51].

Inbred strains are a useful tool for understanding the mode of action of pharmacological agents or toxic substances, and for the identification of molecular pathways that confer susceptibility to therapeutic or adverse events. *In vivo* pharmacogenomic screens have often been carried out in single inbred rodent strains in order to standardize the genotype and limit as many variables as possible. This approach provides mechanistic information in a single genetic background. However, the human situation is far more complex, with diverse genetic backgrounds and uncontrolled environmental effects. Interpretation and translation of data generated from a single inbred strain may lead to generalizations and the inability to distinguish small biologically relevant changes from background variation [2,52]. The use of various models will allow assessment of gene–environment interaction and determination of the genetic causes of inter-individual variability.

5.7 APPLICATIONS OF *IN VIVO* MODELS IN PHARMACOGENOMIC RESEARCH

Over the past decade, pharmacogenomics has benefited greatly from technological advances that have allowed the sequencing of mammalian genomes and the development of multilevel "systems" (high-throughput "polyomic") approaches. For example, gene expression arrays have allowed a global view of transcriptional changes in response to pharmacological intervention or xenobiotic exposure [53,54]. These arrays, in combination with panels of inbred rodent models, have determined whole-genome response signatures, identifying gene clusters functioning within the same biological pathways.

5.7.1 Quantitative Trait Mapping

When embarking on an investigation into the genetic basis of disease, drug response, or toxicity, it is not known where in the genome the causative genes may lie. One approach is to search for evidence that the responsible genes are located within specific chromosomal regions, by detecting correlations between the phenotype and genotype at loci throughout the genome. This approach is known as quantitative trait mapping, and it allows the detection of chromosomal regions harboring genes responsible for disease or drug susceptibility [3,55,56]. Standard intercrosses between inbred strains of rats or mice have been used successfully for mapping quantitative trait loci (QTL). This is carried out by cross-breeding two inbred strains to produce a genetically segregating population that demonstrates quantitative variation in the phenotype of interest. The quantitative phenotype is measured in each individual, and the individuals are then genotyped at a number of genetic markers. A statistical assessment (e.g., the logarithm of the odds (LOD) score) is calculated between the phenotypic values and each marker.

QTL mapping has been used to investigate numerous disease and drug susceptibility traits. One example is the mapping of a mouse limbic seizure susceptibility locus on mouse chromosome 10 by investigating susceptibility to pilocarpine, a muscarinic cholinergic agonist that models human temporal lobe epilepsy [57]. Another example is the identification of QTLs for alcohol-related behavioral traits [58,59].

5.7.2 Identification of eQTLs

Gene expression QTL (eQTL) mapping is a relatively recent technique for assessing associations between transcript expression and genotype in order to identify regulatory locations within the genome [60]. eQTL analysis has shown that genetic regulation of gene expression is a key contributor to population diversity, and it has made an important impact on our understanding of the effects that genetic polymorphisms have on interindividual, tissue-specific variability in physiological processes [61]. The investigation of genes underlying eQTL hotspots allows investigators to determine regulatory candidates of transcription. eQTLs may act in *cis* (locally) or *trans* (at a distance) to a gene. Although *trans* effects have been shown to be globally important, their individual effects are small and therefore their genome-wide identification remains a challenge. eQTL mapping is hypothesis-generating with regard to transcriptional regulation and the discovery of gene networks or pathways of drug susceptibility. However, it can also be used to narrow down a list of genes for further investigation.

While genome-wide association studies (GWAS) in humans tend to focus almost exclusively on statistical evidence (often giving lesser weight to biological plausibility), animal models provide an opportunity to investigate the functional consequences of causal alleles. The production and use of transgenic models allow the investigation of individual or multiple clusters of candidate genes on the phenotype of interest. The increasing complexity of the transgenic lines available, including knockout, knockin, conditional (i.e., the ability to switch a gene on and off), and tissue/organ-specific, allows detailed examination of "causal" variants.

5.8 RODENT MODELS IN PHARMACOGENOMICS

This section provides a snapshot of cases in which animal models have been used successfully in pharmacogenomic strategies.

5.8.1 Bipolar Disorders

Like many neuropsychiatric disorders, bipolar disease is a complex polygenic disorder with variable penetrance, which has made unraveling its underlying genetics particularly difficult [62]. Illness-causing mutations of bipolar disease are often detrimental to reproductive potential and therefore are evolutionarily selected against. This results in minor frequency alleles that require very large sample sizes to achieve statistical significance. Bipolar disorder is further complicated by phenotypic heterogeneity and is influenced by environmental factors [63].

The combination of both human clinical studies and animal research has been crucial to the progress made in understanding the genetic and neurobiological basis of bipolar disorders. The integration of gene expression data from rodent models with human genetic linkage/association data, together with findings from human postmortem brain tissue, has allowed translational validation of findings, resulting in the identification of candidate genes, pathways, and mechanisms for subsequent targeted, hypothesis-driven research [64].

Some of the behavioral signs and symptoms of bipolar disorder can be achieved by methamphetamine administration in rodents, resulting in manic features in the activation phase and depressive features in the withdrawal phase with a single dose [65]. More complex psychosis is achieved with administration of multiple methamphetamine doses [66].

Previous microarray studies by Niculescu et al. [65], in a mouse model, where methamphetamine was used as a stimulant and co-treated with the mood stabilizer valproate, identified a prioritized list of bipolar candidate genes. Results from this pharmacogenomic approach in mice were cross-validated with data generated from human bipolar genetic studies using a strategy termed "convergent functional genomics." Niculescu et al. utilized this translational pharmacogenomic approach as a tool for tagging genes that have pathophysiological relevance. These findings were recently extended to include additional GWAS datasets, and recent evidence from other lines of work, including human post-mortem data, providing the most comprehensive integration of genetics and functional genomics in the field of bipolar disorders [64,67]. Through these studies, six top candidate genes (ARNTL, MBP, BDNF, NRG1, RORB, and DISC1) were identified that merit future follow-up as either novel or existing therapeutic targets.

Moreover, this important study has allowed a panel of single-nucleotide polymorphisms (SNPs) to be generated based on the top identified candidate genes. Based on this panel, a genetic risk prediction score (GRPS) for bipolar disorder was developed and used in an independent cohort to demonstrate that the GRPS differentiates between patients with bipolar disorder and normal controls in both European American and African American populations.

5.8.2 Anxiety Disorders

Niculescu et al. [68] used a similar pharmacogenomic approach to identify candidate genes as well as potential blood biomarkers for anxiety disorders. In this study they used a mouse model treated with the anxiogenic drug yohimbine and the anti-anxiety drug diazepam to examine gene expression changes in blood and key brain regions for anxiety (prefrontal cortex, amygdala, and hippocampus) [68]. The microarray results were analyzed using a convergent functional genomics (CFG) approach, which integrated the researchers' findings with published human and animal model data. This translational strategy of cross-matching and prioritizing findings identified top candidate genes (including FOS, GABBR1, NR4A2, DRD1, ADORA2A, QKI, RGS2, PTGDS, HSPA1B, DYNLL2, CCKBR, and DBP), brain-blood biomarkers (such as FOS, QKI, and HSPA1B), pathways (such as cAMP signaling), and mechanisms for anxiety disorders.

Importantly, this study demonstrated a major overlap between top candidate genes for anxiety and candidate genes for schizophrenia and bipolar disorder, as well as alcoholism previously identified by the same group. This suggests a possible molecular basis for the frequently observed clinical co-morbidity and interdependence between anxiety and other major psychiatric disorders. Notably, PDE10A and TAC1 are at the overlap of all three major psychiatric domains and may be of major interest for drug development [69–71].

In terms of pharmacotherapy and drug development, some of the candidate genes identified in this study encode for proteins that are modulated by existing pharmacological agents, which may suggest future avenues for rational polypharmacy using currently available agents.

5.8.3 Pharmacogenomics Using Xenograft Mouse Models of Acute Leukemia

Acute myeloid leukemia (AL) is a heterogeneous malignant pathology with a poor global prognosis. The disease occurs when hematopoietic progenitor cells undergo malignant transformation into primitive, undifferentiated cells with abnormal longevity [72]. These cells proliferate abnormally, replacing normal bone marrow tissue and hematopoietic cells. Despite significant therapeutic advances in AL, a large number of patients relapse and frequently present with drug-resistant and refractory leukemia. Due to genetic therapeutic options with particular focus on targeted therapies.

Xenograft mouse models are widely used to identify and test new therapeutic targets and strategies for acute leukemia [72,73]. These models are generated in immunodeficient mice (i.e, NOD/SCID (NS) and NS IL2Rgamma (NSG)), which allow the engraftment of human blast cells via injection into the tail vein or retro-orbital sinus. Prior conditioning by irradiation is often used to improve the engraftment rate. However, it has been shown that lifespan is prolonged and vascular injury is reduced in nonirradiated NGS mice [74]. NS or NSG mice injected with primary or leukemic cell lines strongly display the fundamental biological characteristics of the original human disease. Disease progression is assessed by immunophenotyping of human cells in mouse peripheral blood by flow cytometry analysis and/or by histopathology of bone marrow, spleen, and other organs after sacrifice. Immunocompromised mouse models are valuable tools for understanding the effects of individual drugs or drug combinations at the molecular level, and they provide an important opportunity for pharmacogenomic analysis of AL [72].

One example of evaluation of the xenograft mouse model for pharmacogenomic assessment in AL is the study by Bhadri et al. (2011) [75]. Glucocorticoids (GCs), such as prednisolone and dexamethasone, are critical components of multi-agent chemotherapy protocols used in the treatment of acute lymphoblastic leukemia (ALL) [76] because of their ability to induce apoptosis in lymphoid cells. In the study by Bhadri et al. [75], the NOD/SCID xenograft mouse model was evaluated as a tool to investigate glucocorticoid-induced gene expression. Cells from a glucocorticoid-sensitive xenograft derived from a child with B-cell precursor ALL were inoculated into NOD/SCID mice. When highly engrafted (>70% 8–10 weeks post-transplantation), the mice were randomized into groups receiving either dexamethasone (15 mg/kg by intraperitoneal injection) or vehicle control. Leukemia cells were harvested from mice spleens at 0, 8, 24 or 48 h, and gene expression was analyzed on Illumina WG-6_V3 chips.

The 8 hour dexamethasone-treated time point demonstrated the highest number of significantly differentially expressed genes, with fewer observed at the 24 and 48 h time points and with minimal changes seen across the time-matched controls. Dexamethasone treatment identified 173 upregulated genes and 25 downregulated genes. Upregulated genes included those known to be involved in GC response (TSC22D3, SOCS1, FKBP5, MAPK5, ZBTB16). At the later time points, most of the differentially expressed genes are involved in cell cycle progression and apoptotic pathways, consistent with growth arrest. When compared to publicly available datasets of glucocorticoid-induced gene expression from an *in vitro* cell line study and from an *in vivo* study of patients with ALL, expression changes in the 8 h xenograft samples showed a similar response to patients treated with glucocorticoids. It was concluded that the NOD/SCID ALL xenograft mouse model provides biologically relevant insights into glucocorticoid-induced gene expression, in a consistent, reproducible, and clinically relevant model system.

5.8.4 Pharmacogenomics of Pain and Analgesia

Pain severity ratings and the analgesic dosing requirements of patients with apparently similar pain conditions may differ considerably between individuals. Genetic, environmental, and epigenetic factors have been implicated in interindividual differences in pain sensitivity [77]. Evidence of a genetic component of pain modulation comes from rodent heritability studies as well as from familial aggregation and twin studies in humans. More than 350 candidate pain genes have been identified as potentially contributing to heritable differences in pain sensitivity [77].

A study by Tapocik et al. (2009) [54] identified candidate genes and networks specifically associated with analgesic tolerance to morphine. The researchers used a congenic strategy, whereby only those genes specifically correlated with the trait of interest were identified as candidate genes significantly influencing the development of morphine tolerance. Two inbred mouse strains (C57BL/6J (B6), DBA2/J (D2)) and two reciprocal congenic strains (B6D2, D2B6), in which the proximal region of chromosome 10 (Chr10) was introgressed into opposing genetic backgrounds, were used in a behavior genetics approach. Therapeutically relevant doses of morphine were administered to the mice, and acute analgesic sensitivity to morphine was measured by the hot-plate test. Tolerance to morphine developed most rapidly in B6 mice followed by the B6D2 congenic strain (on the D2 genetic background); it did not develop in the D2 mice and developed only slightly in the D2B6 congenic strain (on the B6 genetic background). This indicated a strong influence of the proximal region of Chr10 in the development of tolerance.

In conjunction with behavioral studies, large-scale gene expression profiling was performed using brain sections from the prefrontal cortex (PFC), ventral striatum (VS), temporal lobe (TL), and periaqueductal grey (PAG) regions. The microarray comprised the National Institutes on Aging KO set of 15,247 cDNA clones [78] and the National Institutes of Mental Health Brain Molecular Anatomy Project set of 11,136 cDNA clones (representing ~21,000 distinct mouse genes). Canonical pathway and network analyses were used to identify significant candidate genes for predisposition and tolerance. Canonical pathway analysis of the morphine-responsive genes suggests a significant role for neuroadaptive processes in response to chronic morphine administration (i.e., LTP, axonal guidance, ephrin, neuroregulin pathways).

Network analysis was used by Tapocik et al. [54] in identifying highly interconnected genes to provide insight into molecular events that predispose or are causally related to the development of tolerance. A potential gene network was identified in the PAG in which 19 of 34 genes identified were strongly associated with tolerance. Eleven of the network genes were found to reside in quantitative trait loci previously associated with morphine-related behaviors, whereas seven were predictive of tolerance (morphine-naive condition). Overall, the genes modified by chronic morphine administration showed a strong presence in canonical pathways representative of neuroadaptation. A potentially significant role for micro-RNA and epigenetic mechanisms in response to chronic administration of pharmacologically relevant doses of morphine was highlighted by candidate genes Dicer and H19. Expression of H19 appears to decrease in the genotypes that become tolerant to morphine and increase in those that do not.

5.8.5 Pharmacogenomics and PAH

Pulmonary arterial hypertension (PAH) is a heterogeneous disorder, and it is widely recognized that different clinical groups of PAH exhibit different degrees of responsiveness to the same therapeutic agents. For example, PAH associated with scleroderma and other inflammatory disorders has a generally poorer response than idiopathic PAH [79]. A proof-of-concept study was carried out by Said et al. in 2011 [80] demonstrating that therapeutic responses in different PAH models are reflected by differential underlying gene expression patterns. This study showed that PAH may benefit from targeted therapies based on unique genomic features. In the study, two experimental models of the disease were investigated: mice with deletion of the vasoactive intestinal peptide gene (VIP$^{-/-}$) and Sprague Dawley rats injected with monocrotaline (MCT). These two models express comparable phenotypic features, with the MCT model showing generally more severe PAH [81,82]. Both models were treated with the same drug, vasoactive intestinal peptide (VIP), and the phenotypic and genotypic responses in each were analyzed. Microarray expression profiling in lungs from VIP KO mice demonstrated uniform expression of a single set of gene alterations, comprising upregulation of vasoconstrictor/proliferative genes, downregulation of vasodilator/ antiproliferative genes, and upregulation of inflammatory genes. Treatment of the mice with VIP corrected the key features of PAH, including elevated pulmonary artery pressure, vascular remodeling, right ventricular hypertrophy, and lung inflammation, and it reversed the genetic expression changes toward normal values. However, in the MCT model VIP treatment incompletely corrected the phenotypic changes and was only partially successful in reversing the genotypic abnormalities. MCT genotypic analysis showed two distinct sets of expression alterations: one similar to that in VIP$^{-/-}$ mice (i.e., alterations that promote vascular remodeling and inflammation) and the other suggesting an effort to modulate the PAH.

The authors concluded that, despite its limited scope, this proof-of-concept study provided evidence to support the applicability of pharmacogenomics in PAH. The correlation of genomic analysis with drug response in two animal models demonstrated the usefulness of genomic information in selecting appropriate, targeted therapy in PAH and in predicting therapeutic outcome. However, these findings in rodent models need to be validated using lung tissues from human patients with different forms of PAH.

5.8.6 Pharmacodynamic/Pharmacogenomic Modeling of Insulin Resistance

Glucocorticoids are secreted by the adrenal cortex and maintain adequate blood glucose levels in order to satisfy the demands of the vital organs, particularly the brain. Among other functions, glucocorticoids increase the synthesis and activity of enzymes in the gluconeogenic pathway in liver and kidney and also control the release of gluconeogenic substrates from muscle and fat. Corticosteroids (CS), which are synthetic glucocorticoids, are used clinically for a wide variety of conditions that require immune and/or inflammatory modulation. However, therapeutic use of corticosteroids is accompanied by a wide range of adverse effects that include muscle wasting, insulin resistance, and hyperglycemia, caused by the magnification of "glucocorticoid-type" pharmacological actions [83].

Studies by Almon et al. [84–88] investigated CS effects on insulin resistance genes in rat skeletal muscle. Acute and chronic administration of methylprednisolone (MPL) was carried out in adrenalectomized Wistar rats. Adrenalectomy was used in order to eliminate circadian oscillation in CS-responsive genes to facilitate effective data mining. On completion of the respective CS dosing periods, microarray expression profiling was carried out in mRNA extracted from the rat gastrocnemius muscle using Affymetrix gene chips. Using two different dosing regimens in a time series format, the researchers were able to produce kinetic/dynamic models for the regulation of the expression of six genes that contributed to CS-induced insulin resistance in skeletal muscle. These genes were identified as insulin receptor substrate-1 (IRS-1), uncoupling protein 3 (UCP3), pyruvate dehydrogenase kinase isoenzyme 4 (PDK4), fatty acid translocase (FAT), glycerol-3-phosphate acyltransferase (GPAT), and endothelin-1 (ET-1). The kinetic models included both direct effects of CS on the genes as well as indirect effects due to the effects of CS on other transcription factors such as MyoD and SREBP-1c.

Importantly, these studies demonstrated the complexity of the underlying mechanisms of CS-induced insulin resistance and the necessity of developing experimentally testable hypotheses (dynamic models) on a genome-wide scale to begin elucidating the implicated regulatory pathways.

5.9 CONCLUSION

No single study that uses an isolated approach will answer the important questions relevant to the molecular, physiological, and clinical consequences of a particular polymorphism. Indeed, it is the collection of studies, which include *in vitro* investigations, human clinical studies, and relevant animal models, that together provide the information necessary to inform a pharmacogenomic strategy for a given drug.

Animal models have greatly improved our understanding of the cause and progression of human genetic diseases, and they have proven to be a useful tool for discovering targets of therapeutic drugs. Investigation of a variety of models, appropriate computational support, and continued improvements in technology will increase the sensitivity, speed, and accuracy with which differentially expressed genes are identified. Genome sequencing will incrementally enhance the quality of annotated DNA and protein databases, and will facilitate functional characterization of differentially expressed genes.

REFERENCES

[1] Karczewski KJ, Daneshjou R, Altman RB. Chapter 7: Pharmacogenomics. PLoS Comput Biol 2012;8:e1002817.

[2] Liggett SB. Genetically modified mouse models for pharmacogenomic research. Nat Rev Genet 2004;5:657–63.

[3] Delles C, McBride MW, Graham D, Padmanabhan S, Dominiczak AF. Genetics of hypertension: from experimental animals to humans. Biochim Biophys Acta 1802;2010:1299–308.

[4] Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ, Scherer S, et al. Rat genome sequencing project consortium: genome sequence of the brown norway rat yields insights into mammalian evolution. Nature 2004;428:493–521.

[5] Rusyn I, Gatti DM, Wiltshire T, Kleeberger SR, Threadgill DW. Toxicogenetics: population-based testing of drug and chemical safety in mouse models. Pharmacogenomics 2010;11:127–36.

[6] Szatkiewicz JP, Beane GL, Ding Y, Hutchins L, Pardo-Manuel de Villena F, Churchill GA. An imputed genotype resource for the laboratory mouse. Mamm Genome 2008;19:199–208.

[7] Yalcin B, Wong K, Agam A, Goodson M, Keane TM, Gan X, et al. Sequence-based characterization of structural variation in the mouse genome. Nature 2011;477:326–9.

[8] Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 2011;477:289–94.

[9] Shay T, Jojic V, Zuk O, Rothamel K, Puyraimond-Zemmour D, Feng T, et al. ImmGen consortium. Conservation and divergence in the transcriptional programs of the human and mouse immune systems. Proc Natl Acad Sci USA 2013;110:2946–51.

[10] Pinkert CA. Transgenic animal technology: alternatives in genotyping and phenotyping. Comp Med 2003;53:126–39.

[11] Ogilvie MB. Inbreeding, eugenics, and Helen Dean King (1869–1955). J Hist Biol 2007;40:467–507.

[12] Lazar J, Moreno C, Jacob HJ, Kwitek AE. Impact of genomics on research in the rat. Genome Res 2005;15:1717–28.

[13] Dwinell MR, Lazar J, Geurts AM. The emerging role for rat models in gene discovery. Mamm Genome 2011;22:466–75.

[14] Cohen RM, Rezai-Zadeh K, Weitz TM, Rentsendorj A, Gate D, Spivak I, et al. A transgenic alzheimer rat with plaques, tau pathology, behavioral impairment, oligomeric aβ, and frank neuronal loss. J Neurosci 2013;33:6245–56.

[15] Chia R, Achilli F, Festing MF, Fisher EM. The origins and uses of mouse outbred stocks. Nat Genet 2005;37:1181–6.

[16] Klöting I, Nitschke C, van den Brandt J. Impact of genetic profiles on experimental studies: outbred versus wild rats. Toxicol Appl Pharmacol 2003;189:68–71.

[17] Festing MFW. Inbred strains of rats. Behav Genetics 1981;11:5.

[18] Valdar W, Solberg LC, Gauguier D, Burnett S, Klenerman P, Cookson WO, et al. Genome-wide genetic association of complex traits in heterogeneous stock mice. Nat Genet 2006;38:879–87.

[19] Johannesson M, Lopez-Aumatell R, Stridh P, Diez M, Tuncel J, Blázquez G, et al. A resource for the simultaneous high-resolution mapping of multiple quantitative trait loci in rats: the NIH heterogeneous stock. Genome Res 2009;19:150–8.

[20] Baud A, Herman R, Guryev V, Stridh P, Graham D, McBride MW, et al. Rat genome sequencing and mapping consortium. Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. Nat Genet 2013;45:767–75.

[21] The Complex Trait Consortium. The collaborative cross, a community resource for the genetic analysis of complex traits. Nat Genet 2004;36:1133–7.

[22] Iancu OD, Darakjian P, Walter NAR, Malmanger B, Oberbeck D, Belknap J, et al. Genetic diversity and striatal gene networks: focus on the heterogeneous stock-collaborative cross (HS-CC) mouse. BMC Genomics 2010;11:585.

[23] Aylor DL, Valdar W, Foulds-Mathes W, Buus RJ, Verdugo RA, Baric RS, et al. Genetic analysis of complex traits in the emerging collaborative cross. Genome Res 2011;21:1213–22.

[24] Collaborative Cross Consortium. The genome architecture of the collaborative cross. Mouse genetic reference population. Genetics 2012;190:389–401.

[25] Broman KW. The genomes of recombinant inbred lines. Genetics 2005;169:1133–46.

[26] Hubner N, Wallace CA, Zimdahl H, Petretto E, Schulz H, Maciver F, et al. Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. Nat Genet 2005;37:243–53.

[27] Graham D, McBride MW, Brain NJ, Dominiczak AF. Congenic/consomic models of hypertension. Methods Mol Med 2005;108:3–15.

[28] Jeffs B, Negrin CD, Graham D, Clark JS, Anderson NH, Gauguier D, et al. Applicability of a “speed” congenic strategy to dissect blood pressure quantitative trait loci on rat chromosome 2. Hypertension 2000;35:179–87.

[29] Negrín CD, McBride MW, Carswell HV, Graham D, Carr FJ, Clark JS, et al. Reciprocal consomic strains to evaluate y chromosome effects. Hypertension 2001;37:391–7.

[30] Simmons D. The use of animal models in studying genetic disease: transgenesis and induced mutation. Nature Educ 2008;1:1.

[31] Gama Sosa MA, De Gasperi R, Elder GA. Animal transgenesis: an overview. Brain Struct Funct 2010;214:91–109.

[32] Nagy A. Cre recombinase: the universal reagent for genome tailoring. Genesis 2000;26:99–109.

[33] Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, et al. Knockout rats produced using designed zinc finger nucleases. Science 2009;325:433.

[34] Geurts AM, Moreno C. Zinc finger nucleases: new strategies to target the rat genome. Clin Sci (Lond) 2010;119:303–11.

[35] Carroll D. Zinc-finger nucleases as gene therapy agents. Gene Ther 2008;15:1463–8.

[36] Woods IG, Schier AF. Targeted mutagenesis in zebrafish. Nat Biotechnol 2008;26:650–1.

[37] Mashimo T, Kaneko T, Sakuma T, Kobayashi J, Kunihiro Y, Voigt B, et al. Efficient gene targeting by TAL effector nucleases coinjected with exonucleases in zygotes. Sci Rep 2013;3:1253.

[38] Cade L, Reyon D, Hwang WY, Tsai SQ, Patel S, Khayter C, et al. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. Nucleic Acids Res 2012;40:8001–10.

[39] Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC, Phipps EL. Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. Proc Natl Acad Sci USA 1979;76:5818–9.

[40] Russell LB, Bangham JW, Stelzner KF, Hunsicker PR. High frequency of mosaic mutants produced by N-ethyl-N-nitrosourea exposure of mouse zygotes. Proc Natl Acad Sci USA 1988;85:9167–70.

[41] Smits BMG, Mudde JB, van de Belt J, Verheul M, Olivier J, Homberg J, et al. Generation of gene knockouts and mutant models in the laboratory rat by ENU-driven target-selected mutagenesis. Pharmacogenet Genomics 2006;16:159–69.

[42] Reijmers LG, Coats JK, Pletcher MT, Wiltshire T, Tarantino LM, Mayford M. A mutant mouse with a highly specific contextual fear-conditioning deficit found in an N-ethyl-N-nitrosourea (ENU) mutagenesis screen. Learn Mem 2006;13(2):143–9.

[43] Hemann MT, Fridman JS, Zilfou JT, Hernando E, Paddison PJ, Cordon-Cardo C, et al. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes *in vivo*. Nat Genet 2003;33:396–400.

[44] Sliva K, Schnierle BS. Elective gene silencing by viral delivery of short hairpin RNA. Virol J 2010;7:248.

[45] Larson SD, Jackson LN, Chen LA, Rychahou PG, Evers BM. Effectiveness of siRNA uptake in target tissues by various delivery methods. Surgery 2007;142:262–9.

[46] Tanaka T, Oka T, Shimada Y, Umemoto N, Kuroyanagi J, Sakamoto C, et al. Pharmacogenomics of cardiovascular pharmacology: pharmacogenomic network of cardiovascular disease models. J Pharmacol Sci 2008;107:8–14.

[47] de la Cruz N, Bromberg S, Pasko D, Shimoyama M, Twigger S, Chen J, et al. The rat genome database (RGD): developments towards a phenome database. Nucleic Acids Res 2005;33:D485–91.

[48] Festing S, Wilkinson R. The ethics of animal research. Talking point on the use of animals in scientific research. EMBO Rep 2007;8:526–30.

[49] Robinson V, Jennings M. Refinement and reduction in production of genetically modified mice: sixth report of BVAAWF/FRAME/RSPCA/UFAW joint working group on refinement. Lab Anim 2003;37:1–51.

[50] Robinson V. Finding alternatives: an overview of the 3Rs and the use of animals in research. School Sci Rev 2005;87:319.

[51] Aitman TJ, Critser JK, Cuppen E, Dominiczak A, Fernandez-Suarez XM, Flint J, et al. Progress and prospects in rat genetics: a community view. Nat Genet 2008;40:516–22.

[52] Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, et al. Inflammation and host response to injury, large scale collaborative research program. Genomic responses in mouse models poorly mimic human inflammatory diseases. PNAS 2013;110:3507–12.

[53] Rothberg BE. The use of animal models in expression pharmacogenomic analyses. Pharmacogenomics J 2001;1:48–58.

[54] Tapocik JD, Letwin N, Mayo CL, Frank B, Luu T, Achinike O, et al. Identification of candidate genes and gene networks specifically associated with analgesic tolerance to morphine. J Neurosci 2009;29:5295–307.

[55] Dominiczak AF, Negrin DC, Clark JS, Brosnan MJ, McBride MW, Alexander MY. Genes and hypertension: from gene mapping in experimental models to vascular gene transfer strategies. Hypertension 2000;35:164–72.

[56] McBride MW, Charchar FJ, Graham D, Miller WH, Strahorn P, Carr FJ, et al. Functional genomics in rodent models of hypertension. J Physiol 2004;554:56–63.

[57] Winawer MR, Gildersleeve SS, Phillips AG, Rabinowitz D, Palmer AA. Mapping a mouse limbic seizure susceptibility locus on chromosome 10. Epilepsia 2011;52:2076–83.

[58] Crabbe JC, Phillips TJ, Gallaher EJ, Crawshaw LI, Mitchell SR. Common genetic determinants of the ataxic and hypothermic effects of ethanol in BXD/Ty recombinant inbred mice: genetic correlations and quantitative trait loci. J Pharmacol Exp Ther 1996;277:624–32.

[59] Crabbe JC. Quantitative trait locus gene mapping: a new method for locating alcohol response genes. Addict Biol 1996;1:229–35.

[60] Petretto E, Mangion J, Dickens NJ, Cook SA, Kumaran MK, Lu H, et al. Heritability and tissue specificity of expression quantitative trait loci. PLoS Genet 2006;2:e172.

[61] Grieve IC, Dickens NJ, Pravenec M, Kren V, Hubner N, Cook SA, et al. Genome-wide co-expression analysis in multiple tissues. PLoS One 2008;3:e4033.

[62] Need AC, Goldstein DB. Whole genome association studies in complex diseases: where do we stand? Dialogues Clin Neurosci 2010;12:37–46.

[63] Nakatani N, Aburatani H, Nishimura K, Semba J, Yoshikawa T. Comprehensive expression analysis of a rat depression model. Pharmacogenomics J 2004;4:114–26.

[64] Le-Niculescu H, McFarland MJ, Mamidipalli S, Ogden CA, Kuczenski R, Kurian SM, et al. Convergent functional genomics of bipolar disorder: from animal model pharmacogenomics to human genetics and biomarkers. Neurosci Biobehav Rev 2007;31:897–903.

[65] Ogden CA, Rich ME, Schork NJ, Paulus MP, Geyer MA, Lohr JB, et al. Candidate genes, pathways and mechanisms for bipolar (manic–depressive) and related disorders: an expanded convergent functional genomics approach. Mol Psychiatry 2004;9:1007–29.

[66] Sokolov BP, Polesskaya OO, Uhl GR. Mouse brain gene expression changes after acute and chronic amphetamine. J Neurochem 2003;84:244–52.

[67] Patel SD, Le-Niculescu H, Koller DL, Green SD, Lahiri DK, McMahon F, et al. Coming to grips with complex disorders: genetic risk prediction in bipolar disorder using panels of genes identified

through convergent functional genomics. Am J Med Genet B Neuropsychiatr Genet 2010;153B:850–77.

[68] Le-Niculescu H, Balaraman Y, Patel SD, Ayalew M, Gupta J, Kuczenski R, et al. Convergent functional genomics of anxiety disorders: translational identification of genes, biomarkers, pathways and mechanisms. Transl Psychiatry 2011;1:e9.

[69] Charych EI, Jiang LX, Lo F, Sullivan K, Brandon NJ. Interplay of palmitoylation and phosphorylation in the trafficking and localization of phosphodiesterase 10A: implicationsfor the treatment of schizophrenia. J Neurosci 2010;30:9027–37.

[70] Frisch P, Bilkei-Gorzo A, Racz I, Zimmer A. Modulation of the CRH system bysubstance P/NKA in an animal model of depression. Behav Brain Res 2010;213:103–8.

[71] Mathew SJ, Vythilingam M, Murrough JW, Zarate Jr CA, Feder A, Luckenbaugh DA, et al. A selective neurokinin-1 receptor antagonist in chronic PTSD: a randomized, doubleblind, placebo-controlled, proof-of-concept trial. Eur Neuropsychopharmacol 2011;21:221–9.

[72] Guihard S, Peyrouze P, Cheok MH. Pharmacogenomic considerations of xenograft mouse models of acute leukemia. Pharmacogenomics 2012;13:1759–72.

[73] Barrett DM, Seif AE, Carpenito C, Teachey DT, Fish JD, June CH, et al. Non-invasive bioluminescent imaging of primary patient acute lymphoblastic leukemia: a strategy for pre-clinical modelling. Blood 2011;118:e112–7.

[74] Watanabe S, Ohta S, Yajima M, Terashima K, Ito M, Mugishima H, et al. Humanized NOD/SCID/IL2Rgamma(null) mice transplanted with hematopoietic stem cells under nonmyeloablative conditions show prolonged life spans and allow detailed analysis of human immunodeficiency virus type 1 pathogenesis. J Virol 2007;81:13259–64.

[75] Bhadri VA, Cowley MJ, Kaplan W, Trahair TN, Lock RB. Evaluation of the NOD/SCID xenograft model for glucocorticoid-regulated gene expression in childhood B-cell precursor acute lymphoblastic leukemia. BMC Genomics 2011;12:565.

[76] Gaynon PS, Carrel AL. Glucocorticosteroid therapy in childhood acute lymphoblastic leukemia. Adv Exp Med Biol 1999;457:593–605.

[77] Smith MT, Muralidharan A. Pharmacogenomics of pain and analgesia. Rev Clin Genet 2012;82:321–30.

[78] Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, Wang X, Grahovac MJ, et al. Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. Proc Natl Acad Sci USA 2000;97:9127–32.

[79] Humbert M, Sitbon O, Simonneau G. Treatment of pulmonary arterial hypertension. N Engl J Med 2004;351:1425–36.

[80] Said SI, Hamidi SA. Pharmacogenomics in pulmonary arterial hypertension: toward a mechanistic, target-based approach to therapy. Pulm Circ 2011;1:383–8.

[81] Said SI, Hamidi SA, Dickman KG, Szema AM, Lyubsky S, Lin RZ, et al. Moderate pulmonary arterial hypertension in male mice lacking the vasoactive intestinal peptide gene. Circulation 2007;115:1260–8.

[82] Schermuly RT, Dony E, Ghofrani HA, Pullamsetti S, Savai R, Roth M, et al. Reversal of experimental pulmonary hypertension by PDGF inhibition. J Clin Invest 2005;115:2811–21.

[83] Yao Z, Hoffman EP, Ghimbovschi S, DuBois DC, Almon RR, Jusko WJ. Pharmacodynamic/Pharmacogenomic modeling of insulin resistance genes in rat muscle after methylprednisolone treatment: exploring regulatory signaling cascades. Gene Regul Syst Biol 2008;2:141–61.

[84] Almon RR, Chen J, Snyder G, DuBois DC, Jusko WJ, Hoffman EP. In vivo multi-tissue corticosteroid microarray time series available online at Public Expression Profile Resource (PEPR.). Pharmacogenomics 2003;4:791–9.

[85] Almon RR, DuBois DC, Piel WH, Jusko WJ. The genomic response of skeletal muscle to methylprednisolone using microarrays: tailoring data mining to the structure of the pharmacogenomic time series. Pharmacogenomics 2004;5:525–52.

[86] Almon RR, Dubois DC, Jin JY, Jusko WJ. Pharmacogenomic responses of rat liver to methylprednisolone: an approach to mining a rich microarray time series. AAPS J 2005;7:E156–94.

[87] Almon RR, DuBois DC, Jin JY, Jusko WJ. Temporal profiling of the transcriptional basis for the development of corticosteroid-induced insulin resistance in rat muscle. J Endocrinol 2005;184:219–32.

[88] Almon RR, Lai W, DuBois DC, Jusko WJ. Corticosteroid regulated genes in rat kidney: mining time series array data. Am J Physiol 2005;289:E870–82.

This page is intentionally left blank

Chapter 6

Human Lymphoblastoid Cell Lines in Pharmacogenomics

Jae-Pil Jeon
Division of Brain Diseases, Center for Biomedical Science, Korea National Institute of Health, Korea Centers for Disease Control and Prevention, Republic of Korea

Chapter Outline

6.1 Overview 89
6.2 Generation of LCLs 90
  6.2.1 What Are LCLs? 90
  6.2.2 EBV-Mediated B-Cell Transformation 90
  6.2.3 LCL Generation in Biobanks 92
  6.2.4 LCL Repositories and Collections 93
6.3 Biological Characteristics of LCLs 95
  6.3.1 EBV Transformation Effects on B Cells 95
  6.3.2 Long-Term Subculture Effects on LCLs 96
6.4 Potentials and Limitations of LCL Applications 97
  6.4.1 Genomic Studies Using LCLs 97
  6.4.2 Pharmacogenomic Studies Using LCLs 100
  6.4.3 Immunological Studies Using LCLs 103
  6.4.4 Stem Cell Research Using LCLs 103
  6.4.5 Factors to Consider in LCL Applications 104
6.5 Future Perspectives 105
Acknowledgments 107
References 107

6.1 OVERVIEW

In the post-genome era, there has been a paradigm shift in human genetics and health sciences from the human genome to humanity's genes. In turn, the focus of research has been from individuals toward populations. Different individuals have different genomes in ethnically diverse populations. In terms of gene–environment interaction, mutations or defects in the same gene do not always result in the same clinical outcomes in humans. Likewise, the same polymorphic changes do not always result in the same phenotypes in human physiology. The development and outcome of diseases depend on an individual's genetic makeup interacting with environments such as life style and environmental exposure.

Modern health sciences need unprecedentedly large-scale samples and health data for understanding disease risk factors and therapeutics as well as translating biomedical research into healthcare. Thus, biobanks and next-generation -omics technologies lead to the era of big data, in which personal genetic and health information will be key elements to facilitate the realization of predictive, personalized, preventive, and participatory medicine.

In the field of pharmacogenomics, large numbers of genetic variants underlying drug toxicity and efficacy need to be investigated at the genome-wide level using large numbers of samples and electronic health records. In such efforts, biobanking and biorepositories have been implemented in many areas of public health science, such as discovery science for drug and biomarker development, clinical trials monitoring the effects of drugs on eventual outcomes, and clinical care for therapy. Furthermore, harmonization and data sharing of biobank resources have become a necessity for achieving sufficient statistical power in biobank-assisted pharmacogenomic studies.

Personalized medicine seeks to dissect individuals' responses to environmental stresses or stimuli. In the last decade, a novel *in vitro* LCL-based assay system was developed as a powerful tool to dissect genetic determinants underlying interindividual variations in drug responses and molecular targets of drugs of interest. Despite the limitations of LCLs, pharmacogenomic studies have utilized the *in vitro* cell model using LCLs, and the model's utility is increasingly extended to various genomics and pharmacogenomics fields. Especially with the advent of next-generation sequencing technology, more and more RNA-seq data will be generated from various panels of LCLs which will facilitate the elucidation of global landscapes of gene

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00006-2
© 2014 Elsevier Inc. All rights reserved.

expression in response to environmental stimuli. Here, LCL applications in pharmacogenomics fields and accompanying methodological issues will be discussed.

6.2 GENERATION OF LCLs

6.2.1 What Are LCLs?

LCLs are EBV-transformed lymphoblastoid B-cell lines, which can be easily established by EBV infection or derived spontaneously from peripheral blood B lymphocytes. It is easy to generate LCLs by EBV infection from a small quantity (1−5 ml) of whole blood of healthy donors or patients; it is also easy to maintain LCL cells in continuous subculture. Thus, LCLs are used as an unlimited resource for various genetics and immunologic studies.

Normal B lymphocytes have a limited lifespan *in vivo* and *in vitro* as other human differentiated cells, possibly because of telomere shortening (telomerase activity decrease). However, spontaneous cellular immortalization has been observed in human somatic cells from patients with genetic disorders [1–4]. Spontaneous LCLs can be generated in tissue culture from human patient tissues, including bone marrow cells [5], tonsils [6], and lymphocytes [7]. The transformation efficiency of spontaneous lymphoblastoid cell lines (SP-LCLs) was increased by tumor promoter treatment [8], radiation [9], and sometimes in cyclosporine-containing medium from EBV-seropositive patients [10].

Generally the EBV-mediated B-cell transformation takes 3–4 weeks from an initial EBV infection to the freezing of propagated cells for permanent storage. It is considered a failure of transformation if EBV-infected B lymphocytes fail to form cell clumps by 8 weeks. In addition, when the LCL is capable of proliferating up to the passage number of 160 during the continuous long-term subculture, it is considered to be terminally immortalized.

6.2.2 EBV-Mediated B-Cell Transformation

6.2.2.1 LCL Generation Protocols

The Epstein-Barr virus (EBV) was first described in 1964 as an etiological agent of Burkitt's lymphoma, which was initially found in the African malaria belt [11]. Several years later, it was demonstrated that EBV can induce *in vitro* transformation of leukocytes [12,13]. EBV belongs to the herpes virus family and infects more than 95% of the adult population; the majority of primary infections occur in early childhood and then become generally subclinical for life. However, in patients with reduced immunity, EBV infection often results in diseases, including post-transplant lymphoproliferative disease (PTLD) in organ transplant patients or oral hairy leukoplakia with nonmalignant hyperplastic lesion of epithelial cells.

As in Burkitt's lymphoma, EBV is implicated in many cancers, including nasopharyngeal carcinoma, Hodgkin's disease, and lymphoproliferative disease with impaired T-cell immunity. EBV has a dual cell tropism mediated through the interaction of viral gp350/220 and cellular CD21/CR2 (also the receptor for the C3d component of complement) on B lymphocytes [14]. The EBV genome encodes over 85 open reading frames in the genome, making up a linear dsDNA molecule of about 172 kb. EBV-positive or -negative cell lines have been established from EBV-related lymphoma tissues [15,16]. These EBV-positive lymphoma cell lines contain a different number of EBV genomes (e.g., 2 copies/cell in Namalwa, 50 copies/cell in Raji, 152 copies/cell in Daudi), compared to EBV-negative lymphoma cell lines with no EBV genome (e.g., Ramos, DG75).

6.2.2.2 Transformation Efficiency

Since it was shown that EBV can selectively transform B lymphocytes to produce immortalized polyclonal B-cell lines *in vitro* [17], many efforts have been made not only to understand the mechanism of EBV-mediated B-cell transformation but also to increase the transformation efficiency of lymphocytes that may be important specifically for population-based biobanking. In gradient fractionations of B lymphocyte subpopulations, the target subpopulations of EBV infection are high-density cells representing small resting B lymphocytes [2].

In contrast, the low-density fractions are enriched for large activated B lymphocytes that are resistant to EBV infection. EBV penetrates all B-cell populations, including resting and activated cells, but only the high-density populations are transformed. Only small portions (~20%) of *in vitro* EBV-infected peripheral blood B lymphocytes exhibit EBNA expression, which is the earliest detectable EBV-encoded nuclear antigen after infection [18]. After EBV infection, their cells are variable in size, morphology, and Ig expression. LCLs become oligoclonal or monoclonal because of the outgrowth of selected cell clones during the EBV transformation and immortalization [19].

During EBV transformation, some factors affect transformation efficiency. Factors contributing to the high success rate of EBV transformation include: processing in a sterile environment against microbial contamination; protecting cross-contamination of cell lines; using high-potency viral supernatant of sufficient titer; using good-quality blood samples (e.g., sufficient number of lymphocytes more than 10^6, not contaminated with red blood cells) [20]. Regression of transformed cultures and the delay between collection and processing of blood samples may also contribute to transformation failure. LCL generation methods are well documented elsewhere [21,22]. Various methods and modifications have been described to generate LCLs (Table 6.1).

TABLE 6.1 Major Achievements and Modifications in Protocols for LCL Generation

Keywords	Achievements/Modificaitons	References
EBV discovery	EBV discovered in Burkitt's lymphoma tissue	[11]
Spontaneous LCL	Spontaneous LCL generation first reported	[23]
EBV transformation	EBV used for leukocyte transformation	[12]
B95.8 strain	EBV producer cell lines generated from marmoset leukocytes	[24]
Selective B-cell transformation	EBV selectively transforms B lymphocytes	[17]
Cyclosporine	T cell immunosuppressant increases transformation efficiency	[6]
Frozen lymphocytes and PHA	Frozen lymphocytes used for EBV infection in the presence of PHA	[25]
Frozen lymphocytes and cyclosporine	Frozen lymphocytes used for EBV infection in the presence of cyclosporine	[26]
Small volume of frozen blood cells	LCL generation tested in the context of biobanking	[5]
FK506	FK506 (T-cell immunosuppressant) promotes proliferation of EBV-infected B cells	[27]

Besides controlling these experimental environments, some treatments (e.g., cyclosporine, phytohemagglutinin, CD19 treatment) have been used to increase the efficiency of EBV-mediated B-cell transformation or to shorten the period of immortalization. For example, cyclosporin A, currently known as an immunosuppressive drug, facilitates *in vitro* outgrowth of EBV-infected B lymphocytes over T lymphocytes to increase EBV transformation efficiency [6,28]. This effect is most probably due to inhibition of the immune T lymphocytes, which is implicated in the spontaneous regression of EBV-infected lymphocytes. As most adults are EBV-seropositive and have regulatory T cells, EBV infection can elicit cytotoxic T lymphocytes in cultures that destroy the EBV-infected cells.

Besides functional inactivation of T cells, transformation efficiency can be enhanced by the enrichment of B cells for EBV infection using magnetic beads that conjugate with anti-CD19 antibody specific to B cells [5]. It is known that cyclosporine A inhibits T cell activation by blocking gene expression of cytokines (e.g., IL-2, IL-4), which results in immunosuppressive activity [29]. Recently, another T cell immunosuppressant, FK506 (Tacrolimus), was used for quicker and more reliable EBV-mediated B-cell transformation with a combination of high titers of EBV virus [27]. In general, one can see clumps of cells by light microscope about a week after EBV infection, but their presence is not a specific indicator of successful EBV transformation. Notably, a recent study showed that successful EBV transformation can be monitored by the presence of CD23^{hi}CD58^{+} cells as early as three days post–EBV infection [30].

Phytohemagglutinin (PHA) is a mitogenic agent that has long been used as a red-cell agglutinating agent, removing red blood cells from leukocyte culture. Studies identified PHA as a T cell mitogen that can trigger cell division transforming rapidly to blast cells. The blastogenesis of T lymphocytes leads to cell death before the cytotoxic T cell response is induced against EBV infection. In fact, pretreatment of cord blood leukocytes with phytohemagglutinin results in a 50% increase in EBV-mediated transformability [18]. PHA is less effective than cyclosporin A when blood is from EBV-seropositive donors [21]. Most HapMap LCLs are established by EBV transformation of peripheral blood mononuclear cells using phytohemagglutinin. For example, most LCLs from Coriell Cell Repositories were established using standard EBV transformation protocols that included mononuclear cell separation by Histopaque-1077 Hybrid Max gradient centrifugation in order to obtain peripheral blood mononuclear cells (PMBCs), and induced lymphocyte divisions by phytohemagglutinin as a mitogen [31].

When a new method for EBV-mediated transformation was developed for cryopreserved lymphocytes, fresh blood samples were no longer necessary for EBV transformation as a starting material [25,32]. This method was primarily based on Gusella's protocol [6], but it was modified to include additional steps such as Ficoll-gradient centrifugation of buffy coats prior to cryopreservation and phytohemagglutinin treatment on EBV infection. The Pressman group's protocol [26] included the following steps, which are different from the Louie and King protocol:

1. Direct separation on a density gradient without first isolating the buffy coat to save time
2. Washing the lymphocytes only twice and at lower speeds to be less harsh on the cells
3. Using cyclosporine a in the transformation step rather than PHA-P
4. Using a controlled-rate cell freezer (cyromed model 700a) for a more consistent and accurate rate of freezing to give a better rate of recovery of frozen cells

6.2.2.3 EBV Producer Cell Lines

The most commonly used EBV producer cell line is the B95-8 cell line derived from EBV-infected leukocytes of marmoset monkey. In humans, EBV-transformed B cells or immortalized lymphoblastoid cell lines restrict replication of the virus and release very few extracellular infectious viral particles. But monkey marmoset leukocytes were successfully transformed to release sufficient amount of infectious EBV virions in cultures [24]. In this EBV-transformed marmoset cell line, B95-8, the release of infectious EBV virus particles were increased during the exponential phase of cell growth, but decreased once the cells reached the stationary phase.

Although the B95-8 cell line (ATCC CRL-1612) is the most popular EBV producer cell line, it is no longer distributed commercially without a permit because it is derived from a primate species listed as endangered by the U.S. Fish and Wildlife Service. As an alternative, EBV-containing supernatants from the B95-8 cell line are distributed by ATCC as VR-1492 (http://www.atcc.com). Thus, the culture supernatant of the B95-8 cell line has been used as an EBV viral source for EBV-mediated B-cell transformations in almost all LCL repositories. Besides the B95-8 cell line, other EBV producer cell lines are occasionally used for EBV-mediated B-cell transformation. These cell lines were established from EBV-related lymphoma cells. For example, the Akata, P3HR1, and BJAB cell lines were derived from human Burkitt's lymphoma cells, and the M-ABA cell line was established from a nasopharyngeal carcinoma.

6.2.3 LCL Generation in Biobanks

6.2.3.1 Current Biobanks

A biobank is generally defined as a collection of human biological samples and associated information organized in a systematic way for research purposes. The traditional concept of DNA and serum banks has evolved to current biobanks in the genetic epidemiological studies, boosting population-based biobanks for healthy donors. In contrast, tissue repositories for cancer research were advanced for the pathology-based human biobanks to foster disease-based biobanks for patients. Although human biobanks have to consider additional, complicated ethical issues regarding human rights and dignity, they are considered to be one of the biological resource centers (BRCs) for which global organizations have developed guidelines dealing with operational and quality issues, biosecurity, collaborations, and so forth. Biobanking and biorepositories have been implemented in many areas of public health, such as drug and biomarker development, clinical trials monitoring of the effects of drugs on eventual outcomes, and clinical care for therapy.

6.2.3.2 Blood Cell Banking for LCL Generation

In general, when population-based biobanks collect blood-derived samples, whole blood samples are stored in liquid nitrogen tanks for EBV transformation for future purposes. To generate LCLs, starting points for EBV-mediated B-cell transformation can be anywhere in the process of peripheral blood fractionation, such as whole blood itself, buffy coats, peripheral blood mononuclear cells (PBMCs) isolated by Ficoll-gradient centrifugation, or purified CD19(+) B cells (Figure 6.1). It was reported that a small quantity of frozen blood cells (as little as 500–800 ul) was successfully used to generated LCLs with a transformation success rate of over 95% [5] (http://www.hpacultures.org.uk/services/gss/EBVtransformation.jsp).

First of all, in order for whole blood to be used for direct EBV transformation, it should be anti-coagulated in citrate- (ACD-) or EDTA-containing tubes. In most cases, the ACD vacutainer is recommended for drawing peripheral blood for later use in EBV transformation. If not used immediately, the blood sample needs to be cryo-preserved in 10% DMSO in liquid nitrogen. However, successful EBV transformation was also achieved when residual clinical blood samples were placed on the bench for a considerable time before EBV infection [33]. Interestingly, the U.K. Biobank's pilot study showed that when blood samples were maintained at room temperature for variable periods of 24 h, 1 or 2 weeks before cryo-preservation and subsequent positive B-cell selection using the Miltenyi autoMACS cell separation station (Miltenyi Biotec), EBV transformation could be achieved at a success rate of 92%, 67%, or 38% for blood samples, respectively [5]. This pilot study provides population-based biobanks with a more cost-effective method based on small volumes (0.5 ml) of cryo-preserved whole blood. Thus, whole-blood samples could be successfully cryo-preserved for the future generations of LCLs to be selectively transformed for future needs.

6.2.3.3 Quality Control of LCLs

Because EBV producer cell lines are used as an EBV source, human lymphoblastoid cell lines are occasionally

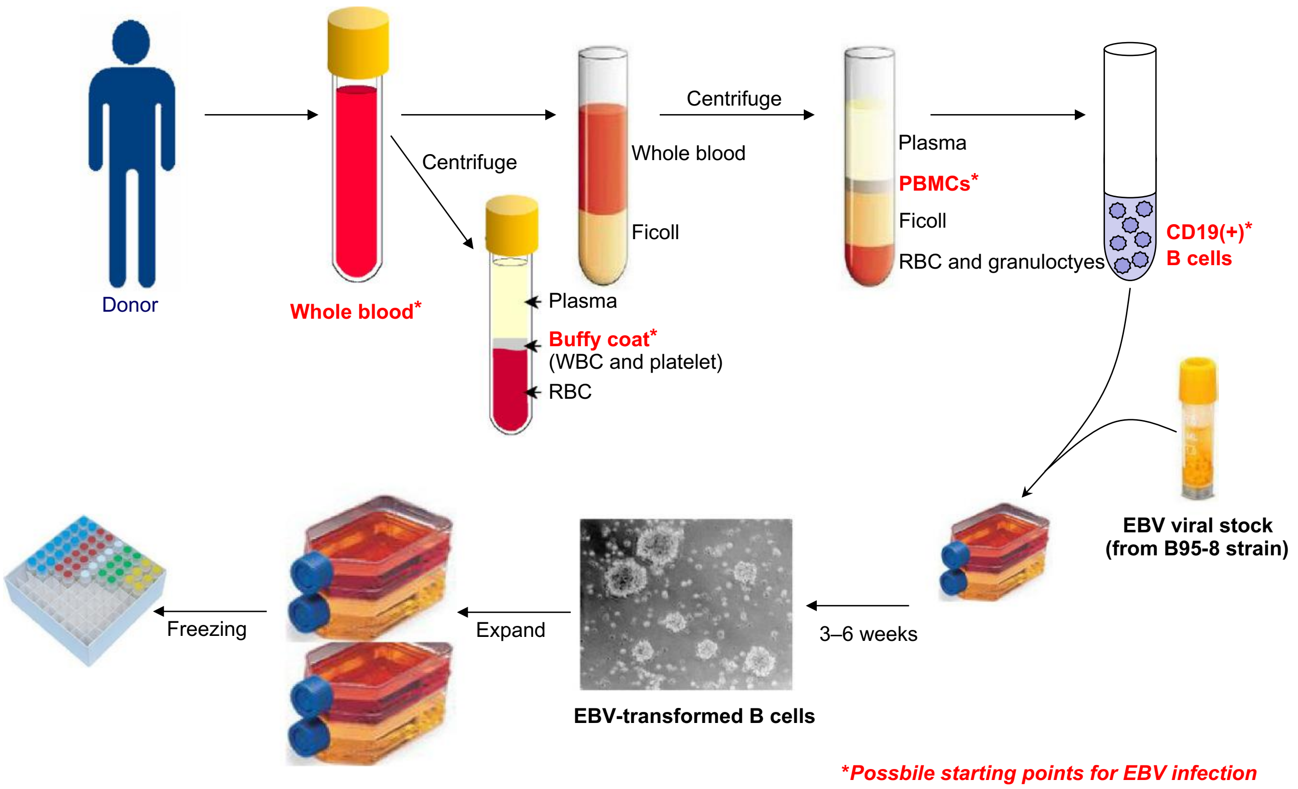


FIGURE 6.1 LCL generation workflow.

contaminated with them during LCL generation. Several standard techniques—DNA fingerprinting by STR (short tandem repeats) analysis, SNP genotyping analysis, human leukocyte antigen (HLA) typing, and the like—should be used for cell line authentication. In addition, large-scale cell repositories often encounter cell line cross-contamination and misidentification. Studies have highlighted the number of cell lines that have been cross-contaminated with another cell line (often older, more well-established cancerous cells) [34,35].

The International Cell Line Authentication Committee (ICLAC) has maintained a database of cross-contaminated or misidentified cell lines [36]. For example, it is known that some 15–20% of cell lines in current use may be cross-contaminated or mislabeled in routine culture [37]. According to MacLeod et al. [35], 18% of 252 cell lines donated to the bank were misidentified or contaminated. Interestingly Drexler et al. [38], reported that some EBV-positive B lymphoblastoid cell lines was misclassified as myeloma cell lines (e.g., IM-9, GM1312, ARH-77) or Hodgkin's lymphoma cell lines (e.g., L-591, RB, RPMI-6666). Besides cell authentication, other quality control measures should be taken for microorganism contamination (e.g., bacterial and mycoplasma) during LCL generation and maintenance. LCLs should be handled using aseptic techniques under the Biosafety Level 2 (BSL-2).

6.2.4 LCL Repositories and Collections

Because the value of LCLs was demonstrated in many applications of genomics and pharmacogenomics and immunological studies, large LCL collections have been established from various populations with different ethnic origins. The most well-known LCL collections are the CEPH pedigrees, the International HapMap populations, and the Human Variation Panel populations, providing a vast public resource of genotype and gene expression data.

6.2.4.1 CEPH Collections

The CEPH LCLs (http://www.cephb.fr) consists of 1064 LCLs from 1050 individuals in 51 world populations, including 30 trios from the international HapMap CEU samples (northern and western European ancestry from Utah). Most CEPH LCLs were derived from Caucasian individuals of European multigenerational families, representing general, but not disease-targeted, populations. Since the article by Cann et al. [39] on the HGDP-CEP Human Genome Diversity Cell Line Panel, three standardized subsets of the original diversity panel have been recommended for future population genetic studies [40]. For example, the first subset (denoted H1048) contains 1048 samples of the original HGDP-CEPH panel, excluding one member of each duplicate pair and two individuals with extremely atypical genotypes for their population. The second subset (denoted H971) excludes 77 individuals from the H1048 in order to avoid including two individuals with a first-degree relationship (parent/offspring or full sibling). The third subset (denoted H952) excludes additional 19 individuals from the subset H971 so that the H952 contains no two individuals with a second-degree relationship (half-sibling, avuncular, or grandparent/grandchild).

The CEPH LCLs have been extensively genotyped for genetic variants such as SNPs and CNVs; these data are publicly available (http://snpdata.cshl.edu/population_studies/linkage_maps//). Most importantly, CEPH cell panels provide publicly available baseline gene expression data of the families, as in the Gene Expression Omnibus of CEPH pedigrees (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1485) and CEU RNA-Seq data (http://jungle.unige.ch/rnaseq_CEU60). These CEPH LCLs can be obtained from the CEPH collection through the Coriell Institute for Medical Research (http://locus.umdnj.edu/nigms/ceph/ceph.html).

6.2.4.2 HapMap Collections

The International HapMap LCL collections are one of the most studied LCL collections (http://www.hapmap.org). The HapMap Phase I/II comprises 270 LCL cell lines from Caucasian groups (CEU) plus three different ethnic groups (YRI, Yoruban people in Ibadan, Nigeria; CHB, Han Chinese in Beijing; JPN, Japanese in Tokyo). These LCL panels were derived from 30 CEPH trios from the Caucasian CEU population, 30 trios from the Nigerian YRI population, 45 unrelated people from the Han Chinese CHB population, and 45 unrelated people from the Japanese JPT population. For the HapMap Phase III project, an additional 1031 LCL cell lines were derived from seven other populations (ASW, people of African ancestry in the Southwestern U.S.; GIH, Gujarati Indians in Houston, Texas; LWK, Luhya people in Webuye, Kenya; MEX, people of Mexican ancestry in Los Angeles, California; CHD, Chinese in Metropolitan Denver, Colorado; MKK, Maasai in Kinyawa, Kenya; and TSI, Toscani in Italy) [41]. These International HapMap LCLs cells are also available from the Coriell Institute for Medical Research (http://ccr.coriell.org).

6.2.4.3 NBK Collections

In the last decade, a large number of biobanks have been established to study human genome diversity and gene–environment interaction in the context of personalized medicine. These biobanks have collected various human biological samples, including blood-derived and urine samples. Some biobanks, like the U.K. Biobank, store blood-derived cell samples (e.g., whole blood, buffy coats, or peripheral blood mononuclear cells) for future applications (e.g., LCL generation). In contrast, the National Biobank of Korea (NBK), which is Korea's biggest population-based biobank, separated peripheral blood mononuclear cells from ACD-mediated anti-coagulated whole blood and subsequently conducted EBV transformation in the presence of cyclosporine A. The NBK is responsible for biobanking of biological samples from various population-based cohort studies for the Korean Genome and Epidemiology Study (KoGES); for further information, see http://www.nih.go.kr/NIH/eng/contents/NihEngContentView.jsp?cid=19822&menuIds=HOME004-MNU0710-MNU1028-MNU1029.

So far, the NBK has established large-scale LCL collections of approximately 15,000 LCLs from various cohort studies (http://kbn.cdc.go.kr). These collections comprise about 11,000 LCLs derived from community-based urban and rural cohort participants, about 2500 LCLs from twin cohort participants, and about 1500 LCLs from elderly cohort participants. Among these NBK LCLs, the Korean HapMap LCLs were derived from 90 unrelated healthy Korean individuals of an equal sex ratio and age (40–69 years at the time of baseline study).

Among the Korean HapMap LCLs, 20 LCL strains were randomly selected for study of their biological characteristics and genomic changes during LCL generation and maintenance [42]. This special collection comprised 17 immortal LCL strains that were able to proliferate more than the passage number of 160 (p160) during continuous long-term subculture, and three nonimmortal LCL strains that stopped proliferation around the passage number 40 (p40), at which point LCLs undergo cellular crisis. Most of the NBK LCL samples are linkable to health information as well as to biological samples (e.g., peripheral-blood DNA, serum, plasma, urine) from the same donors. Moreover, some donors have been genotyped for genetic variations (e.g., SNPs, CNVs) in genome-wide association studies. A large collection of NBK LCLs will provide a good resource for genome-wide expression studies as well as pharmacogenomic discovery science, which requires a sufficient sample size to represent ethnic diversity.

6.2.4.4 NLGIP Collections

The National Laboratory for the Genetics of Israeli Populations (NLGIP) provides research communities with 1673 LCLs from various Jewish and Arab populations in Israel (http://www.tau.ac.il/medicine/NLGIP/catalog.htm). All NLGIP donors were over 18 years old and Israeli citizens. LCL cells and matching DNA samples from the NLGIP have been used to study the polymorphism of drug-metabolizing enzymes in Israeli populations, including CYP2C9, CYP2C19, and CYP2D6. Recently, these NLGIP LCLs were extensively utilized in pharmacogenomic studies of LCL gene expression profiles.

6.2.4.5 Other LCL Collections

Many LCL collections can be found in the PubMed literature and websites. The MACS (Multicenter AIDS Cohort Study) holds EBV-transformed B-cell lines and PBMC pellets in addition to plasma and serum samples from over 7000 participants (http://www.statepi.jhsph.edu/macs/macs.html). MACS is an ongoing prospective study of the natural and treated histories of HIV-1 infection in homosexual and bisexual men, conducted at study sites located in multicenters in Baltimore, Chicago, Pittsburgh, and Los Angeles [43,44].

The International Histocompatibility Working Group (IHWG) also provides a comprehensive LCL inventory which is maintained by the Research Cell Bank (RCB) located at Fred Hutchinson Cancer Research Center (http://www.ihwg.org/). The core inventory consists of more than 1000 cell lines from the tenth, twelfth, and thirteenth IHWS, from HLA heterozygous and homozygous donors, from selected families, and from individuals of diverse population groups. It maintains stocks of purified DNA from LCLs and DNA reference panels that provide an extensive array of HLA and HLA-related genetic variants. Collectively, these cell lines represent more than 375 unique HLA-A, B, C, DRB1, DQA1, DQB1, DPA1, and DPB1 alleles.

The Avon Longitudinal Study of Parents and Children (ALSPC) is a long-term health research project also known as Children of the 90s (http://www.alspac.bris.ac.uk). It enrolled more than 14,000 mothers during pregnancy in 1991 and 1992, and has followed the health and development of their children over two decades. Biological samples, including blood, urine, hair, nails, saliva, and placenta, have been collected from children and their parents since early pregnancy. Until now, approximately 15,000 LCLs have been generated from children (n>7100), mothers (n>5700), and partners (n>1900). In addition, approximately 22,000 samples of genomic DNA from blood, LCLs, and mouthwash samples can be provided for research projects as outlined in ALSPAC policy. ALSPAC provides a rich resource for the study of the environmental and genetic factors that affect health and development. Thus it aims to inform policy and practices that will provide a better life for future generations.

The Cedars-Sinai Medical Center published its prolonged experiences of EBV B-cell transformation in 1991. They established about 2900 LCLs by EBV transformation, about 600 of which were from frozen lymphocytes. Success rates using freshly isolated versus previously frozen lymphocytes are approximately 97% [26].

6.3 BIOLOGICAL CHARACTERISTICS OF LCLs

The following subsections discuss the effects of EBV transformation and long-term subculture on biological characteristics of LCLs.

6.3.1 EBV Transformation Effects on B Cells

In general, EBV-transformed B lymphocytes are highly variable in shape and tend to make cell clumps in stationary cultures [45]. However, cell clustering is not a specific indicator of EBV-mediated growth transformation. Ficoll/Hypaque-gradient centrifugation is used to separate the interface containing peripheral mononuclear cells from the pellet and lower gradients containing the majority of red cells and mature granulocytes. Generally EBV-transformed B lymphocytes grow in suspension with immunoglobulin synthesis, but EBV infection of B lymphocytes occasionally generates unusual lymphoblastoid cell lines that grow uniform and round with no detectable synthesis of immunoglobulin and lysozyme [45].

EBV infection mimics the effect of B-cell activation signals such that EBV-infected B cells exhibit higher proliferation activity than resting B cells [46]. The B-cell activation method was applied to shorten the time required for EBV-mediated B-cell transformation. For example, Hur et al. [47] reported that EBV-infected B cells were activated to accelerate the cellular transformation process when treated with anti-human CD19 and anti-human IgM monoclonal antibodies. Generally, the transformation takes about 4 weeks, and cell clumps of increased cell size can be seen at 1–3 weeks after EBV infection. This CD19-immunoglogulin co-ligation method was able to synergistically shorten the time for generating LCLs without changes in typical phenotypes of EBV transformation such as cell surface molecule expression, which accelerates cell clump formation within at least 14 days. On the other hand, Oh et al. [48] reported that the use of EBV viral stocks prepared from TPA-activated B95-8 cells not only shortened the time of LCL generation but also increased the success rate of immortalization.

6.3.1.1 Immunophenotypes

EBV-transformed B cells have been found to be mainly positive for CD38 (like germinal center B cells). IgD and CD38 cell surface molecules have been widely used for B-cell phenotyping in association with B-cell development. EBV-infected B cells undergo changes in immunophenotypes of IgD and CD38 cell surface molecules. That is, the resting B cells represent the IgD(+)CD38(−) cells, while the EBV-infected B cells represent the IgD(−)CD38(+) cells. This is similar to the germinal center B-cell phenotype. On the other hand, colligation of the membrane protein CD19 with the antigen receptor of B lymphocytes decreases the threshold for antigen receptor–dependent stimulation so that the B cell can be activated in response to low concentrations of antigens [49].

In early cellular events of EBV infection, two distinct subpopulations of CD23(+) B cells exhibit different destinations for proliferation of EBV-infected B cells. Indeed, in one study the expression pattern $CD23^{hi}CD58(+)IL6(-)$

predicted the identity of infected cells destined for proliferation as early as three days after exposure to EBV. In contrast, another subpopulation marked with CD24^{lo}CD58(+) IL(+) failed to proliferate despite the fact that LMP1 and EBNA2 were expressed in individual CD23hiCD58+ and CD23loCD58(+) cells, demonstrating that reduced levels of LMP1 do not explain the lack of CD23loCD58(+) cell proliferation. This result suggests that B-cell subpopulations are differentially susceptible to EBV-mediated proliferation [30].

6.3.1.2 Gene Expression Changes

Microarray analyses of 10 LCLs and corresponding donors' B cells reveal that EBV-infected B cells exhibit gene expression changes involved in cell proliferation, cell cycle, and immune response, compared to resting B cells [7]. Specifically, 232 genes, including *STMN1*, *RAB9*, *RAB27A*, and *MYBL2* were found to be upregulated in LCLs compared to resting B cells, while 111 genes, including *BACH1*, *BACH2*, *RAB37*, *JUNB*, and *VAV3* were downregulated in LCLs. Resting B cells represented mostly cell cycle–arrested cells in the G0 + G1 phases, while EBV-infected B cells exhibited more proliferating cells in the S and G2 + M phases. Thus, differentially expressed genes between resting and EBV-infected B cells may explain changes in cell cycle profiles in the EBV-infected B-cell transformation.

6.3.2 Long-Term Subculture Effects on LCLs

6.3.2.1 Growth Properties

Once EBV-transformed B cells overcome a proliferative cellular crisis in a middle passage number (around 40), they become capable of continuously proliferating up to a passage number of 160–180, at which point the LCLs are considered to be terminally immortalized [50]. Many studies support biological differences between LCLs at early passages after EBV infection and LCLs at late passages after long-term subculture. To some extent, it is convenient to conceptually discriminate the early LCLs (eLCLs) at a passage number approximately less than 30–40 from terminally immortalized LCLs (tLCLs) at a passage number of more than 160. It is known that the cellular immortalization of LCLs is affected by many factors, including culture conditions, EBV viral activities, and host genomic backgrounds.

In our hands, most of LCL strains (17 out of 20) could proliferate up to a passage number of 160 (p160) in the continuous long-term subculture in which culture medium was exchanged with fresh complete medium at every passage. The remaining three LCL strains stopped proliferation around an average passage number of 41. These nonimmortal LCLs exhibited relatively lower copy numbers of mitochondrial DNA and EBV genome than terminally immortalized LCLs, suggesting that reduced mitochondrial activity may be indicative of diminished EBV-mediated proliferative activity. Thus, in order for LCLs to be immortalized and maintained during long-term subculture, they require sustained EBV viral activities, including EBV viral gene expression and viral genome copy numbers as well as telomerase activity [51,52].

6.3.2.2 Gene Expression Phenotypes

Long-term LCL subculture–induced gene expression changes in NF-kB pathway- and carcinogenesis-related genes that may be involved in the completion of cellular immortalization and EBV-induced tumorigenesis [53]. Indeed, microarray analysis showed that approximately 500–1500 genes per LCL strain were differentially expressed between eLCLs and tLCLs. Of these differentially expressed genes, *CD38* and *PTPN13* were downregulated and upregulated in tLCLs, respectively, compared to eLCLs among all tested 17 LCL strains.

By contrast, 13 genes (*CD180*, *FCRL5*, *GPR160*, *HERC5*, *IFIT1*, *OAS3*, *RASGRP3*, *RFTS7H*, *TC2N*, *TCL1A*, *XAF1*, *ZNF382*- and miR-146a) were downregulated in 16 tLCLs out of 17 tested LCL strains, while *PRKCH* was upregulated in 16 tLCLs. It is known that *PTPN13*, *HERC5*, and miR-146a are associated with NF-kB pathway while *XAF1*, *TCL1A*, and *CD38* are associated with carcinogenesis. In addition, miR-146a is upregulated by LMP1 via NF-kB in lymphocytes. Thus, the NF-kB pathway plays an important role in prolonged LCL subculture. Differentially expressed genes between eLCLs and tLCLs would provide not only molecular signatures required for LCL immortalization but also potential quality control markers related to LCL passage numbers during long-term subculture.

6.3.2.3 miRNA Expression Phenotypes

It is known that miRNAs regulate the expression of their target genes at the levels of transcription and translation, contributing to global gene expression regulatory machineries in cell proliferation, differentiation, and apoptosis. To extend LCL applications to studies using miRNA, it is necessary to understand changes in miRNA expression profiles between eLCLs and tLCLs during LCL generation and maintenance. Using long-term subculture collections of 17 matched pairs of early- and late-passage LCLs, miRNA microarray analysis showed that 46–133 miRNAs per LCL strain were differentially expressed with more than a two-fold change between eLCLs and tLCLs [54]. In particular, miR-99a, miR-125b, and miR-1296 were upregulated in tLCLs compared to eLCLs, while miR-20b*, miR-28-5p, miR-151-3p, miR-151:9.1, miR-216a, and miR-223* were downregulated in tLCLs. In addition, these miRNA expression profiles were able to discriminate eLCLs from tLCLs, as determined by hierarchical clustering analysis.

Interestingly, when target genes of such differentially expressed miRNAs were putatively predicted by miRNA target prediction tools (e.g., TargetScan and miRDB), the corresponding target genes were found to coincide with differentially expressed genes between eLCLs and tLCLs. For example, *CD38*, *FCRL5*, and *XAF1* were target genes of miR-28-5p while *RASGRP3* and *GPR160* were target genes of miR-99a and miR-125b, respectively. These target genes and miRNAs also exhibited negative correlation of expression patterns, which suggests that these pairs of miRNAs and target genes may play roles in LCL maintenance and cellular immortalization. The differentially expressed miRNAs would provide miRNA signatures associated with terminal immortalization of LCLs rather than the initial stage of EBV-mediated B-cell transformation, and could be a potential marker for cellular immortalization and quality control of long-term culture collections of LCLs.

6.4 POTENTIALS AND LIMITATIONS OF LCL APPLICATIONS

It is important for LCL applications in genetic, pharmacogenomic, and immunological studies to know how well LCLs preserve the genetic information and biological characteristics of primary B cells during LCL generation and maintenance. For this reason, various studies have been conducted to evaluate biological and genomic characteristics of LCLs compared to primary B cells. For example, one recent study showed that gene expression profiles and promoter-methylation profiles were affected by EBV transformation while interindividual variations in gene expression levels of the primary B cells were maintained in the LCLs [55]. Here, the uses and applications of LCLs are exemplified to discuss the potential and the limitations of such applications.

6.4.1 Genomic Studies Using LCLs

6.4.1.1 SNPs

The feasibility of using LCL in genetic studies has been extensively evaluated for genomic structural variations, SNP analysis, expression phenotype studies, and the like. The Multicenter AIDS Cohort Study (MACS) studied fidelity of SNP genotype calls in DNA samples that were obtained from paired samples of PBMCs and LCLs from 16 individuals of European American ancestry. They showed that genotypic fidelity increased with more stringent SNP calling quality control. For example, the average mismatch rate was 0.12% for SNPs with a population call rate of at least 95%, and 0.03% for SNPs with a call rate of at least 99%. This work indicated that genotypic discrepancies found in PBMC and LCL pairs were likely to have no apparent gross genotypic errors in the process of EBV transformation, suggesting that mismatch rates of SNP genotypes are minimal with stringent quality control [43].

In addition, SNP fidelity was estimated on the long-term subculture panel of the NKB LCL collection between early and late passage of LCLs during long-term subculture [56]. Genotype calls from the Affymetrix 500K SNP array set were used to estimate genotype concordance between PBMCs and LCLs using identify-by-state (IBS) analysis. The mean pairwise IBS distance of original SNPs was approximately 0.995 between PBMCs and different passage numbers of LCLs, indicating that LCL DNA is a reliable source for genotyping with SNP chip platforms.

However, loss of heterozygosity (LOH) was observed on the loci showing the highest rates of genotype discordance with PBMCs; such LOH abundance increased with higher passage numbers of LCLs. This result suggests that LOH may be a major cause of genotype errors for late-passage LCLs. Substantial genotypic errors were observed mostly in late-passage LCLs (tLCLs), but not in early-passage LCLs (eLCLs, <40 passages). Indeed, overall mean genotype concordance between PBMCs and LCLs was 0.996 for the SNPs that are most frequently used in GWAS. It is therefore likely that LCLs are a reliable DNA source for genetic studies if the DNA sample is harvested from the early-passage LCLs.

Recently, NGS technology was used to study genomic differences between blood and LCL DNA. Londin et al. [31] reported that, on whole-exome sequencing of a tetrad family, an average of 19,354 genetic variants (17,672 SNPs and 1682 indels) were identified per sample in the coding sequences of peripheral-blood mononuclear cells and LCLs from the same donor. They found approximately 99.82% concordance in DNA sequences between PBMCs and LCLs, leaving a de novo mutation rate of 0.18%. Study findings suggest that LCLs are an appropriate representation of the genetic material of the donor but the researchers recommend that PBMC or early-passage LCLs be used.

Linkage analysis using family-based data cannot be applied to infectious diseases such as in susceptibility studies because of the rarity of multi-case family diseases such as HIV infection and tuberculosis. However, the use of an *in vitro* novel strategy based on HIV infection of LCLs successfully identified new genetic variants that influence susceptibility to HIV-1 [57]. The conditions for efficient transduction of LCLs with a VSV.G (vesicular stomatitis virus G protein) -pseudotyped HIV-1-based vector that expresses the GFP reporter gene in transduced LCLs was established. Interestingly, the researchers observed a significant correlation (Pearson $R^2 = 0.56$, $p = 0.007$) between the levels of transduction of CD4+ T cells and LCLs for the same donors, suggesting that LCLs reflect CD4+ T cell behavior. Through a quantitative genome-wide linkage analysis, a region on HSA8q24.3 (marker rs1398296) was identified as the genomic locus contributing the variation in cellular susceptibility to HIV. This genetic variant was also associated with accelerated disease progression in one of two cohorts of HIV-1-infected patients. This suggests that genetic analysis of *in vitro* cellular phenotypes would facilitate the discovery of susceptibility loci to infectious agents.

6.4.1.2 CNVs

In many studies, the genome-wide SNP assay has been used for analyses of copy number variations in LCLs. In one study, for example, a HapMap collection (n = 270 LCLs) was used to generate the comprehensive genome-wide map of copy number of variations (CNVs) in the human genome. This study exploited two technology platforms of the Affymetrix GeneChip Human Mapping 500K array and the CGH Whole Genome TilePath array to generate CNV information from the HapMap collection, consisting of 30 parent-offspring trios of the Yoruba from Nigeria (YRI), 30 parent-offspring trios of European descent from Utah (CEU), 45 unrelated Japanese from Tokyo (JPT), and 45 unrelated Han Chinese from Beijing (CHB). A total of 1447 copy number variable regions (CNVRs) covering 360 megabases—912% of the genome) were identified in diverse populations. Moreover, it was estimated that LCL-specific genomic errors accounted for less than 0.5% of observed copy number variations (CNVs), implying that LCLs are robust sources of genomic DNA for CNV studies [58].

Simon-Sanchez et al. [59] used LCL DNA for a genome-wide SNP assay conducted to detect structural genomic variations in LCLs (n = 270) from neurologically normal healthy donors. They assayed 408,804 SNPs for genotyping using the Infinium Human-1 Genotyping BeadChip (containing 109,365 gene-centric SNPs) and the HumanHap300 Genotyping BeadChip (containing 317,511 SNPs) with 18,072 SNPs in common between the two arrays. The researchers observed extended homozygosity (contiguous tracts >5 Mb) in 9.5% (26/272) of LCLs and 340 structural genomic alterations in 66.9% (182/272) of LCLs. Such structural genomic variations were confirmed in further examination of DNA extracted directly from the blood of 30 of these subjects. This study showed that LCLs have few or only minor effects on genomic structural variation and are unlikely to confound the results of a genome-wide case-control study.

On the other hand, a recent study using LCLs showed the successful discovery of copy number variations by analyzing genome-wide SNP chip data [60]. In initial screening steps, Jeon et al. used SNP array data from the Affymetrix 50K SNP array and the Affymetrix SNP 5.0 array for the Korean HapMap LCLs (n = 90), and then validated CNV regions by quantitative multiplex PCR of short fluorescent fragment (QMPSF). This approach enabled identification of DNA copy number variations at the LEPR gene locus encoding the first two exons of *LEPR* and *LEPROT* genes in LCLs from the Korean population. In addition, this work revealed an association between a structural variation at the LEPR gene locus and complex metabolic traits and the risk of type 2 diabetes.

A recent study reported results of an analysis of the whole-genome sequencing data (40X) for a pair of LCLs and whole blood from the same donor [61]. It determined with high confidence that 99.2% of the genome were identical with no reproducible changes in indels or structural variations, including chromosomal rearrangements and copy number variations. Specifically, when inspected ploidy from genome coverage files in windows of 2kb, the LCL exhibited a decreased copy number in only 4 regions (3 haploid regions, 1 deletion) that contain 4 genes (*KIAA0125*, *PRAME*, *ZNF280A*, *ZNF280B*) and 1 pseudogene (*ADAM6*), compared to the genomic DNA from PBMCs.

6.4.1.3 LCL Transcriptomics

Several gene expression studies showed that LCLs can be used as a potential source of a molecular screen for diseases or environmental stresses. For example, DNA microarray analyses of 116 LCLs from individuals with idiopathic autism identified circadian rhythm–related genes that are differentially expressed in the most severely affected ASD group, providing a clue to circadian rhythm dysfunction in severe autism [62].

On the other hand, another study analyzed transcriptomic profiles of LCLs to identify irradiation-responsive genes and characterize their temporal expression profiles. In this study, CEPH LCLs from 10 unrelated individuals were irradiated at 2 different doses, 3 Gy and 10 Gy, and then harvested at various time points, 0, 1, 2, 6, 12, and 24 hr, after irradiation exposure before pooled RNA samples from all 10 individuals were used for microarray experiments. This genome-wide transcriptional survey of irradiated LCLs demonstrated that LCLs are a very useful model in elucidating the complex transcriptional processes involved in response to DNA damage following irradiation exposure [63].

LCL expression phenotypes (transcript expression levels) can be used to identify potential target genes associated with clinical traits. Correlation studies were conducted to show the utility of LCL expression phenotypes in association studies. Indeed, clinical trait–associated genes were identified from regression analysis using eLCL expression phenotypes and quantitative clinical data from the corresponding donors. For example, gene expression levels of *APOC1* or *NAT12* in eLCLs were significantly associated with 120 min postprandial glucose level in corresponding donors with the positive or negative relationship, respectively, while body mass index (BMI) was positively or negatively correlated to *SERPINB12* or *ACP2*, respectively. Among these significant correlations, the three highest coefficients of determination (R^2) were observed in correlations between CRP (C-reactive protein) levels and *TMSL8* ($R^2 = 0.798$), between BMI and *SERPINB12* ($R^2 = 0.722$), and between 120 min postprandial insulin levels and *ERRF11* ($R^2 = 0.713$). These clinical trait–associated expression phenotypes will facilitate the discovery of novel candidate target genes of diseases or clinical symptoms. For example, it is possible that *APOC1* and *SERPINB12* may be candidate target genes for diabetes and obesity.

However, correlations of expression phenotypes with clinical traits in eLCLs were not replicated in tLCLs, suggesting that the utilization of expression phenotypes should consider LCL passage numbers. Indeed, differential cytotoxicity of anti-cancer agents was observed in pre- and postimmortal LCLs [64]. The postimmortal LCLs undergo cellular immortalization processes such as overcoming the proliferation crisis, increased telomerase activity, abnormal karyotypes, p53 mutation, and downregulation of p16/Rb, mimicking lymphoma cells.

In contrast, the preimmortal LCLs have normal diploid karyotypes and undergo telomere shortening, representing characteristics of actively proliferating normal lymphoblasts. Cytotoxicity tests using pre- and postimmortal LCLs showed that the postimmortal LCLs were significantly more resistant to cytotoxicity of DNA-modifying agents (e.g., camptothecin, etoposide, bleomycin, fluorouracil, thioguanine, melphalan, actinomycin D) while pre- and postimmortal LCLs showed similar levels of cytotoxicity against DNA-nonmodifying agents (colchicines, paclitaxel, vincristine, methotrexate). This result suggests that pairs of pre- and postimmortal LCLs provide a good model for assessing and screening anticancer drugs. That is, if a certain chemical exhibits stronger cytotoxicity in preimmortal LCLs, it is possibly a DNA-modifying agent and may have stronger adverse effects on actively proliferating normal cells. In contrast, if an agent has a markedly stronger cytotoxicity in postimmortal LCLs, it is possibly an ideal anticancer drug because it possibly has weaker side effects on actively proliferating normal cells.

6.4.1.4 *Expression Phenotype Studies (eQTL)*

In contrast to CNV and SNP, the utility of mRNA expression of LCLs is controversial for eQTL analysis. Choy et al. showed that nongenetic factors, such as EBV genome copy number, *in vitro* growth rate, and cellular ATP levels of individual LCLs, were more strongly associated with drug response and mRNA expression phenotypes than genetic factors, such as genotypic variation. Some reports have suggested that LCL mRNA expression data may not be suitable for eQTL studies because of random sampling of LCL clones with monoallelic expression [65,66].

6.4.1.5 *Protein*

Wu et al. [67], in a recent study, showed that protein levels of LCLs are heritable molecular phenotypes that exhibit considerable variation between individuals, populations, and sexes. They used isobaric tag–based quantitative mass spectrometry to determine relative protein levels of 5953 genes in lymphoblastoid cell lines from 95 diverse individuals genotyped in the HapMap project. They also identified protein quantitative trait loci (pQTLs), including variants not detected by previous transcriptome studies. Recently, a protein expression study using LCLs identified the differential expression of DNA repair and cell apoptosis–related proteins, including *RPA32* and p27 in response to radiofrequency radiation [68]. LCL proteomics have been also conducted to discover biomarkers of diseases, including bipolar disorder [69].

6.4.1.6 *DNA Methylation*

DNA methylation is an essential epigenetic mechanism that modulates gene expression and the maintenance of DNA integrity and stability. In order to find genetic variants of dopamine receptor D4 gene (*DRD4*) associated with DNA methylation levels, DNA methylation profiles of 89 LCLs from individuals (from 30 family trios) were interrogated for associations with *cis*-acting genetic variants within the promoter region of dopamine receptor D4 gene [70]. These associations between DNA methylation levels and genetic variants in LCLs were nominally replicated in the postmortem brain tissue that is the *DRD4* target tissue implicated in psychiatric disease phenotypes. This association study provided an alternative approach to assigning a possible functional role to seemingly nonfunctional SNP associations. Recently, Bell et al. [71] reported genome-wide associations between DNA methylation levels with genetic variants and gene expression variations using 77 LCLs from HapMap Yoruba individuals. DNA methylation patterns were used for association analyses with genetic and gene expression variations in LCLs.

6.4.1.7 *RNA-seq*

With the advent of next-generation sequencing (NGS), RNA-seq data have been produced from different sets of LCL panels, including 20 European-ancestry individuals from the Utah pedigrees of Coriell CEPH cell lines [72]; (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29158), 60 LCLs from CEU individuals (CEPH Europeans from Utah) with the Illumina GAII platform [73], and 69 Yoruban individuals (African descent) with the Illumina GAII platform [74]. The RNA-seq analysis of the Utah LCLs detected most of the known genes because only 21 alternatively spliced, multi-exon genes that are not in databases were uncovered at about 40 million sequence reads per sample [72]. In early 2013, RNA-seq data from 1000 genome samples were released. This RNA sequencing dataset of 462 human lymphoblastoid cell line samples from the CEU, FIN, GBR, TSI, and YRI populations from the 1000 Genomes sample collection was created by the Geuvadis consortium (http://www.geuvadis.org/web/geuvadis/home).

6.4.1.8 miRNA

EBV was the first virus found to encode miRNAs. During EBV transformation and cellular immortalization, EBV-infected cells express not only host cellular miRNAs (e.g., miR-155, miR-146a/b, miR-21, miR-103, miR-320a/b/c/d, let-7(a-i)) which account for about 50% of all miRNAs, but also EBV-encoded miRNAs (e.g., miR-BART1-5p/3-5p/4-5p, miR-BHRF1-1/1-2-3p/2-5p), which account for 7–22% of the total miRNAs expressed in LCLs [75]. These host and viral miRNAs are involved in the immune and apoptotic response for establishing a latent infection, ultimately contributing to viral oncogenesis. For example, miRNA-155 was dramatically upregulated by EBV latent membrane protein 1 at the initial stage of EBV infection. In contrast, miRNA-155 was not differentially expressed between early-passage LCLs (eLCLs) and terminally immortalized LCLs (tLCLs) of passage number p160, suggesting that miRNA-155 is required to be sustained for LCL maintenance during long-term subculture.

Recent reports showed the possibility that transcriptomic miRNA profiles of LCLs could be utilized as expression phenotypes for association studies with clinical traits or disease conditions. In the association study using miRNA expression profiles of 20 LCLs and their donors' clinical information, it was observed that expression levels of miR-188-3p and miR-296-3p were significantly associated with systolic blood pressure ($R^2 = 0.202$, $p = 0.047$) and HbA1C levels ($R^2 = 0.272$, $p = 0.018$), respectively. These clinical trait–associated miRNAs (e.g., miRNA-188-3p, miRNA-296-3p) may be new molecular candidate targets for studies of clinical trait–related diseases such as metabolic syndromes and diabetes.

6.4.2 Pharmacogenomic Studies Using LCLs

Many studies have been conducted to discover the relationship of drug sensitivity with genetic variants and global gene expression in LCLs (Figure 6.2). The cytotoxicity assays of LCLs represent a good *in vitro* model for finding genetic variants and target genes associated with drug responses

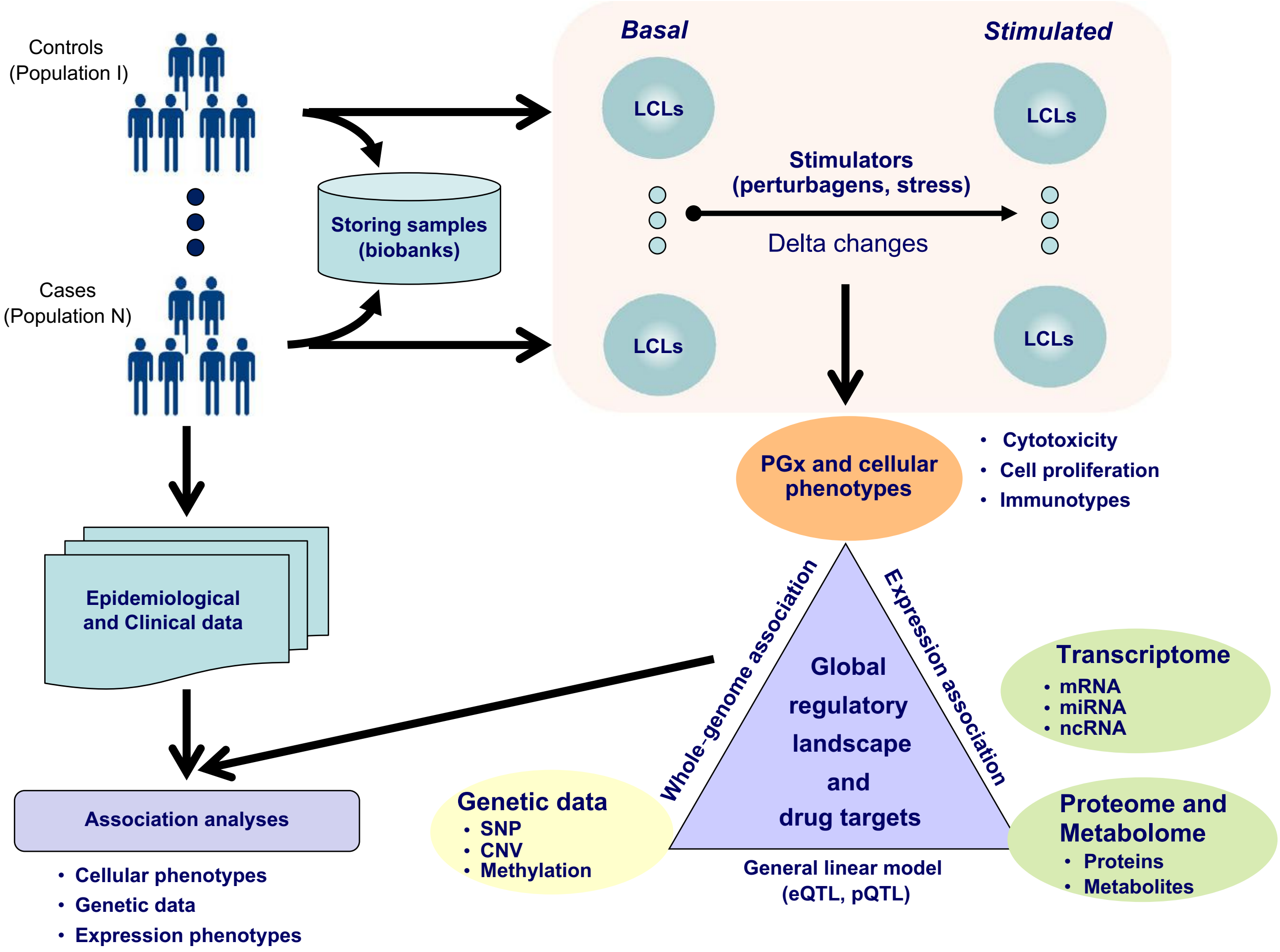

Controls
(Population I)
Cases
(Population N)
Storing samples
(biobanks)
Basal
Stimulated
LCLs
LCLs
LCLs
LCLs
Stimulators
(perturbagens, stress)
Delta changes
PGx and cellular phenotypes
• Cytotoxicity
• Cell proliferation
• Immunotypes
Epidemiological and Clinical data
Association analyses
• Cellular phenotypes
• Genetic data
• Expression phenotypes
Whole-genome association
Expression association
Global regulatory landscape and drug targets
General linear model
(eQTL, pQTL)
Genetic data
• SNP
• CNV
• Methylation
Transcriptome
• mRNA
• miRNA
• ncRNA
Proteome and Metabolome
• Proteins
• Metabolites

FIGURE 6.2 **Applications of LCL-based assays in pharmacogenomic studies.**

in pharmacogenomic research. Most genome-wide association studies employing expression phenotypes of LCLs have used simple association methods using dose-response phenotypes of individual LCLs, where dose—response curves are necessarily nonlinear. Furthermore, many pharmacogenomic studies assume that drug response is a heritable trait, and that differential drug response is defined by one parameter of a complex nonlinear model.

In most pharmacogenomic studies, cell growth inhibition was evaluated for each cell line at different concentrations of a given drug by measuring the concentration required to inhibit 50% cell growth (IC_{50}). Such IC_{50} values of given drugs are first determined for individual LCLs and then used as variables of cellular phenotypes presumably inherent in individual LCLs in association studies. In contrast, the cell proliferation activity of drug-treated cells can also be used as cellular phenotypes for association studies. This method is based on the activity of mitochondrial enzymes that are inactivated shortly after cell death and therefore reflect the number of viable cells after drug exposures.

Cell growth activity is measured either by directly counting viable cells after drug treatment at a certain concentration or by using the colorimetric 2,3-bis(2-methoxy-4-nitro-5-sulfophenyl)-5-([phenylamino] carbony)-2H-tetrazolium hydroxide (XTT) assay in the cell lines treated at drug concentrations for inhibiting 50% of cell growth (IC_{50}). Therefore, either IC_{50} values of drugs for individual LCLs or cell proliferation activity (direct XTT measurements) of drug-treated LCLs can be exploited in association studies such as triangle association and case control, as discussed next.

6.4.2.1 *Triangle Association Model*

In pharmacogenomic studies utilizing LCLs, the most common approach is based on the triangle model that integrates genotype (SNPs), gene expression (transcription profiles), and *in vitro* cytotoxicity (drug responses) data from LCLs. Many genetic components related to drug responses were discovered via the triangle model or the genome-wide genetics of gene expression using LCLs, although such association methods still need to be improved. For example, a recent whole-genome study utilized genotype data and baseline gene expression data of LCLs from three different ethnic populations (CEU/HAPMAPPT01, n = 78; YRI/HAPMAPPT03, n = 87; ASW/HAPMAPPT07, n = 90), where a general linear regression analysis allowed identification of solute carrier transporter (SLC) genes (*SLC31A2*, *SLC43A1*, *SLC35A5* and *SLC41A2*) associated with paclitaxel sensitivity [76].

In addition, eQTLs regulating SLC gene expression were identified using the SCAN database (http://www.scandb.org), and all SNPs were found to be trans-acting eQTLs of their respective SLC gene. That is, rs6935238 was significantly associated both with the baseline gene expression of *SCL31A2* ($P = 8 \times 10 - 5$) and with paclitaxel-induced cytotoxicity in three populations ($P = 1.5 \times 10 - 6$). One eQTL (rs13419945) was significantly associated with cytotoxicity ($P = 4.6 \times 10 - 5$) and baseline gene expression of both *SCL35A5* and *SLC41A2* in LCLs. In these association studies, the use of well-genotyped LCL panels such as the international HapMap take advantage of accumulated genomic data from ethnically diverse populations.

The triangle model has also been used successfully to identify pharmacogenomic candidates associated with the response to many chemotherapeutic agents, including carboplatin [77], gemcitabine and arabinosylcytosin [78], cytarabine arabinoside [79], and tamoxifen [80]. The triangle model has also been applied to radiation pharmacogenomics studies in which integrated analysis was performed with SNPs, basal gene expression, and radiation sensitivity. It is known that genetic inheritance plays an important role in individual variations in radiation response.

Indeed, it was estimated that genetic factors might account for nearly 80% of interindividual radiation response variations. Moreover, peripheral blood lymphocyte radiosensitivity is known to be significantly correlated with clinical response or radiation morbidity. In general, radiation pharmacogenomics studies have identified target genes that might contribute to interindividual variations in radiation response, utilizing genome-wide basal gene expression profiles, genome-wide SNPs, and radiation cytotoxicity obtained from hundreds of irradiated LCLs.

Recently, a radiation cytotoxicity study used the radiation AUC (area under the curve) as an *in vitro* radiation response phenotype from 277 LCLs, including 93 African American (HD100AA), 89 Caucasian American (HD100CAU), and 95 Han-Chinese American (HD100CHI) unrelated healthy subjects [81]. Radiation cytotoxicity differed significantly among the three racial groups. LCLs from Han-Chinese American subjects were more sensitive to radiation than LCLs from Caucasian American subjects. Such radiation AUC for 277 LCLs was used in correlation analysis for the association of genome-wide SNPs and basal expression with radiation cytotoxicity. Subsequent functional validation using siRNA knockdown showed that *C13orf34*, *MAD2L1*, *PLK4*, *TPD52*, and *DEPDC1B* significantly altered radiation sensitivity in cancer cell lines, implying that these genes might be novel biomarkers contributing variations in response to radiation therapy.

6.4.2.2 *Case-Control Design*

The NLGIP LCLs from healthy adult females (n = 80) were phenotyped for drug sensitivity in response to paroxetin, a selective serotonin reuptake inhibitor (SSRI) at a concentration of 10 uM, which represents the mean IC_{50} for growth inhibition [82]. Based on paroxetin sensitivity, 14 "outlier" LCLs were selected for gene expression profiling, which included the most sensitive LCLs (n = 7) and the least sensitive group (n = 7). Of the differentially expressed genes between the two outlier LCL groups (e.g., *ARRB1*, *CCL5*, *DDX60*, *DDX60L*, *ENDOD1*, *ENPP2*, *FLT1*, *GABRA4*, *GAP43*, *MCTP2*, *SPRY2*), the most striking differential gene expression was observed for *CHL1* (cell adhesion molecule with homology to *L1CAM*), which is known to be implicated in correct thalamocortical circuitry, schizophrenia, and autism. This result suggests that these differentially expressed genes, including *CHL1*, may be a promising candidate SSRI response biomarker and that *CHL1* protein may have potential as an antidepressant drug target. This indicates that similar approaches using *in vitro* drug response phenotypes and LCL gene expression profiles can provide a tool to discover drug pathways and drug response–related biomarkers.

In a study [83] undertaken to discover genomic predictors of interindividual sensitivity to a DNA alkylating agent (MNNG), a panel of 24 LCLs was divided into two classes with either high or low sensitivity for a two-class prediction algorithm. Of these cell lines, the four most sensitive and the four least sensitive LCL strains were selected and analyzed to identify differentially expressed genes with significant positive or negative correlation with increasing MNNG sensitivity. They found that increased *MYH* or decreased *C21ORF56* gene expression was correlated with increased alkylating sensitivity, implying that a collective set of another 48 genes, including *MYH* and *C21ORF56*, may be used as predictors of cellular sensitivities to chemotherapeutics. Therefore, a panel of LCLs can be classified into case-control groups based on cellular responses to drugs or environmental toxicants in order to discover genomic predictors or modulators of cellular sensitivities to drugs or compounds.

6.4.2.3 *Clinical Translation of LCL Findings*

Increasing numbers of *in vitro* cell-based pharmacogenomic findings have been successfully replicated in clinical settings. That is, genetic variants identified in the LCL-based model could be used in identifying patients treated with carboplatin with clinical outcome. For example, a recent study demonstrated that LCL-based genome-wide approaches identified carboplatin sensitivity–associated genetic variants that were successfully validated in clinical samples from patients treated with carboplatin [84]. In this genome-wide approach, HapMap LCLs (HAPMAPPT01, CEU, n = 87, for the discovery set) and 52 unrelated CEPH LCLs (for the replication set) were used to generate carboplatin sensitivity variables (IC_{50} values) at concentrations of 0, 10, 40, and 80 uM of carboplatin for 72 h. The SNP genotype and baseline gene expression data of the corresponding LCL samples were downloaded from the International HapMap database (http://www.HapMap.org) and GEO (GenBank Accession No. GSE7761), respectively. This cell-based genome-wide analysis identified 342 SNPs associated with carboplatin-induced cytotoxicity, of which 18 were further found to be associated with baseline gene expression.

One of these SNPs (rs1649942) was replicated in an independent LCL set and is located within the intron of *NRG3*. This SNP genotype exhibited associations with expression levels of 10 genes; the expression levels of 7 of those genes (*ALDH2*, *CRIM1*, *KYNU*, *LOC100131869*, *OAS1*, *RAPGEF5*, and *SLC2A5*) were negatively correlated with carboplatin IC_{50}, while those of 3 genes (*BCR*, *PSTPIP2*, and *SHFM3P1*) were positively correlated with carboplatin IC_{50}. Moreover, it was found that this SNP was associated with progression-free survival (PFS) and overall survival (OS) in phase I analysis of 377 Australian ovarian cancer patients who received at least 4 cycles of carboplatin-based chemotherapy. That is, the risk allele of rs1649942 was associated with a modestly increased risk of disease progression and death following carboplatin-based chemotherapy. However, this finding was not replicated in a larger second-phase evaluation of patient samples, possibly due to differences in clinical definitions and ethnic heterogeneity between the two clinical samples.

In an earlier study, it was shown that LCLs treated with simvastatin (a cholesterol synthesis inhibitor) induced differential expression of full-length and alternative splicing variants of the statin target gene *HMGCR* (3-Hydroxy-3-methylglutaryl coenzyme A reductase) [85]. Statin is known to reduce low-density lipoprotein cholesterol (LDL-C) by inhibiting *HMGCR*, a rate-limiting enzyme for cholesterol synthesis. Such differential induction of *HMGCR* transcripts was associated with *in vivo* response to simvastatin treatment. Greater *in vitro* induction of an alternative splicing variant in LCLs was significantly associated with smaller *in vivo* absolute and percent reductions of total cholesterol, LDL-C, apoB, and triglycerides. Interestingly, *in vivo HMGCRv_1* expression in the absence of statin was not associated with *in vivo* lipid or lipoprotein measurements at baseline.

A major limitation of this approach is that *in vitro* measurements of gene expression were assessed in statin nontarget tissue; statin's primary target tissue is liver. However, there have been many reports that gene expression of lymphocytes or LCLs may reflect nonlymphoid gene expression and may be a marker for hepatic cholesterol metabolism with concordant *HMGCR* regulation between liver and peripheral blood cells [1]. Thus, using both *in vitro* differential isoform induction and *in vivo* drug response facilitates the discovery of new drug targets for improving drug efficacy, and provides a bridge between LCL findings and clinical transitions.

6.4.3 Immunological Studies Using LCLs

6.4.3.1 A Source of Neutralizing Antibodies for Serotherapy

EBV-transformed B cells preserve the immunological characteristics of the original B cells with regard to EBV receptors, complement receptors, surface immunoglobulin, and secretory immunoglobulin [86]. Before EBV transformation, memory B cells carrying appropriate surface immunoglobulin receptors can be preselected by several methods (e.g., fluorochrome-tagged antigens, subcloning, resetting with antigen-coupled erythrocytes). Alternatively, LCL clones producing specific antibodies of interest can be selected from EBV-transformed B cells after EBV infection. These LCLs will carry the appropriate surface immunoglobulin and also secrete the corresponding antibody [87,88]. However, difficulties in cloning LCLs impeded their utilization as a source of antibody-producing cells for immunotherapy.

Traggiai et al. [89] reported an improved method for EBV immortalization of human memory B cells that fully exploits the potential of the human immune response. They found that the TLR9 agonist CpG increased the efficiency of EBV B-cell immortalization by more than 100-fold and facilitated LCL cloning. The LCL clones were stable and secreted antibodies at a high rate (10–50 mg/L in the culture supernatants), which is applicable to antibody production for preclinical studies.

6.4.3.2 A Source of Antigen-Presenting Cells for Immunotherapy

In addition to a source of antibody-producing cells, LCLs that express particular antigens are useful sources of the autologous antigen–presenting cells (APCs) that are often required for restimulating human antigen-specific cytotoxic T lymphocytes (CTLs). Based on the fact that LCLs are potent antigen-presenting cells, their utility was investigated for cancer vaccines against hematological neoplasias and solid tumors [90] and pancreatic cancer [10]. For example, cancer testis antigen- (CTA-) specific T cell response was considerably induced by DNA methyltransferase 1 inhibitor alone or in combination with histone deacetylase inhibitors, which increased CTA genes in LCLs. The autologous SP-LCL was shown to be a good source of antigen-presenting cells (APCs) for cancer immunotherapy against EBV-associated malignancies, including Burkitt's lymphoma, Hodgkin's disease, and nasopharyngeal and gastric carcinomas [91,92].

6.4.4 Stem Cell Research Using LCLs

6.4.4.1 A Source Material of Induced Pluripotent Stem Cells

Recent reports shed light on potential applications of LCLs in generating human-induced pluripotent stem cells (iPSCs) from patient- or healthy donor–derived LCLs via the virus-free episomal method [93,94]. Generally, viral vectors have been used to generate iPSCs where vectors and transgenes are integrated into the genome. This integration problem hampers practical use of iPSCs for clinical applications. Generating LCL-derived iPSCs facilitates the development of human disease models and novel drug therapies. Recently, Choi et al. [93] demonstrated that patient-derived LCLs can be reprogrammed into iPSCs via a virus-free and integration-free method.

Briefly, alpha1-antitrypsin–(AAT-) deficient patient-derived EBV-transformed B cell lines were transfected by EBNA-1/OriP-based episomal vectors encoding a combination of reprogramming genes, including *c-Myc*, *KLF4*, *NANOG*, *LIN28*, and SV40-T antigen. LCL-derived iPSCs exhibited the characteristics of pluripotent stem cells; expressed pluripotency markers, including *TRA-1-60*, *SSEA4*, *OCT4*, and *NANOG*; retained inherited disease-specific donor mutation; and maintained the parental rearranged IgG/V(D)J locus. These patient-specific LCL-iPSCs could differentiate into all three germ layer cell types (endoderm, mesoderm, and ectoderm) as well as directly into mature hepatocyte-like cells (DE cells → hepatic progenitor cells → hepatocytes). These cells also showed no integration of the episomal reprogramming-related transgenes or EBV-related genes, and recapitulated an important disease feature after directed differentiation.

Rajesh et al. [94] also demonstrated that LCLs are reprogrammable to iPSCs via a virus-free and integration-free method using OriP/EBNA-1 plasmid vectors encoding reprogramming genes under feeder-free conditions. They found that a minimal combination of *OCT3/4*, *SOX2*, *NANOG*, and SV40-T antigens was sufficient for reprogramming LCLs, and that iPSCs lost EBNA-1 gene expression as well as the other EBV viral activities after subculture (more than p10 passages) due to the gradual loss of episomal vectors in the absence of drug selection. These LCL-derived iPSCs can be differentiated *in vitro* into hematopoietic, cardiac, neural, and hepatocyte-like lineages.

Thus, LCL-derived iPSCs might provide a useful tool in studying human disease modeling and drug target screening in target-specific cells of interest. LCL-derived iPSCs can take advantage of the easy manipulation of and large quantity of LCL cells. For example, as large collections of LCLs have been generated in population-based or disease-based biobanks in last the decades, LCL cells from patients with rare diseases or from the family pedigree can be easily accessible in many biobanks. More important, many LCL collections have been provided to individual studies or to international genomics or genetics consortia in order to release information about genetic variants and gene expression profiles that can be linked to donor health information, including epidemiological and clinical data. For these reasons, patient-derived

LCL-iPSCs are valuable for disease modeling, especially for rare diseases [95].

6.4.5 Factors to Consider in LCL Applications

6.4.5.1 Nongenetic Confounding Factors

Many LCL application studies showed that interindividual variations in drug-induced LCL growth inhibition are independent of basal growth rate and cell density of each cell line. For example, Morag et al. [82] reported observing no correlations between the degree of paroxetine-induced cell growth inhibition either with cell numbers (basal cell growth rate) 24 h prior to drug exposure or with cell densities immediately prior to drug exposure. This study presented a novel *in vitro* method for distinguishing shared versus distinct drug pathways, demonstrating that drugs with shared pathways exhibit a similar degree of cell growth inhibition while drugs with distinct pathways exhibit different cell growth inhibition profiles. Thus a useful tool is provided for *in vitro* assessment of *in silico*–generated drug pathway predictions and for distinguishing shared versus distinct pathways for compounds of interest.

In addition Fry et al. [83], reported massive differences in sensitivity across a panel of LCLs exposed to a DNA-alkylating agent, N-methyl-N′-nitro-N-nitrosoguanidine (MNNG). LCLs' MNNG sensitivities were not associated with individual differences in growth characteristics (e.g., cell doubling time), so it is likely that interindividual variations in drug response may be attributed to inherent LCL genetic diversity. However, there are many nongenetic confounding factors, as well as biological noise, that may affect *in vitro* assay results for interindividual differences in drug response.

Indeed, one recent study reported biological noises and *in vitro* artifacts that may reduce the ability to detect true relationships and create the potential for nongenetic confounding, causing spurious associations between genetic variants and traits (e.g., expression phenotypes, drug sensitivity) Choy et al. [96] found that the degree of drug response of individual LCLs was similar among three anti-cancer drugs (e.g., 5-fluorouracil, 6-mercaptopurine, and methotrexate), and that the baseline growth rate of individual LCLs was highly correlated to the relative responses to these drugs. Such a correlation could imply the possible influence of an experimental confounder, but it could also suggest a shared genetic mechanism common to these anti-cancer drugs. In fact Morag et al. [82], reported that paired drugs sharing a similar pathway exhibited similar growth inhibition profiles in LCLs, but not drugs with different pathways. For example, good correlations were observed in LCLs for drug responses among antidepressants (e.g., paroxetine, imipramine, fluvoxamine, amitriptyline), among antipsychotics (e.g., clozapine, haloperidol), and among steroid drugs (e.g., dexamethasone, prednisolone, beclomethasone). In contrast, no correlation was found for drug responses in LCLs between antidepressants (e.g., paroxetine) and anticancer drugs (e.g., 6-mercaptopurine, resveratrol, 5-fluorouracil, methotrexate) or between anticancer drugs (e.g., 5-fluorouracil, methotrexate, 6-mercaptopurine).

In utilizing LCLs in pharmacogenomic studies, one should pay careful attention to the many nongenetic confounding factors that may be introduced somewhere in the processes of LCL generation and maintenance as well as in the measurements of cellular phenotypes (Figure 6.3). Biological noises and *in vitro* artifacts are possibly due to the random choice of B-cell subpopulations for EBV-mediated transformation and subsequent clonal expansion of EBV-infected B cells, cell passage numbers, batch effects (including day-to-day variations in all trait measurements and cell culture conditions), and different EBV viral loads per cell, as well as baseline growth rate, ATP levels, and the measurements used to assess drug responses and mRNA expression phenotypes. Therefore, in order to rule out any influence of these potential confounders, additional validation of the assay should be conducted to establish the intra- and interday reproducibility of cellular phenotype measurements.

6.4.5.2 Replication and Functional Validation

Besides the nongenetic confounding factors discussed in the previous section, EBV transformation may affect interindividual variations in cellular phenotypes, including mRNA expression profiles and LCL drug sensitivities of LCLs. EBV-infected B cells undergo a certain degree of biological and genomic changes compared to primary B cells, and they become established immortal cell lines during long-term subculture. EBV transformation effects may vary not only in batch-to-batch LCL generation from the same donor but also in diverse cell lines from different donors. Such EBV transformation may often cause the LCL chromosomal instability and cellular changes that contribute to the possibility of false-positive results in further association studies. EBV transformation effects as well as many potential *in vitro* confounders (batch effects, baseline growth rate, cell passage numbers, etc.) can be compromised by functional validation experiments of such initial findings from LCLs. Thus studies using LCLs need to be functionally validated in other assay systems such as siRNA knockdown experiments in LCLs, or in different cancer cell lines such as the NCI-60 tumor cell line panel. They also need to be further replicated in relevant tissues and in clinical populations.

6.4.5.3 Advantages and Limitations

In vitro model systems using LCLs have revolutionized pharmacogenomic discovery science for personalized medicine.

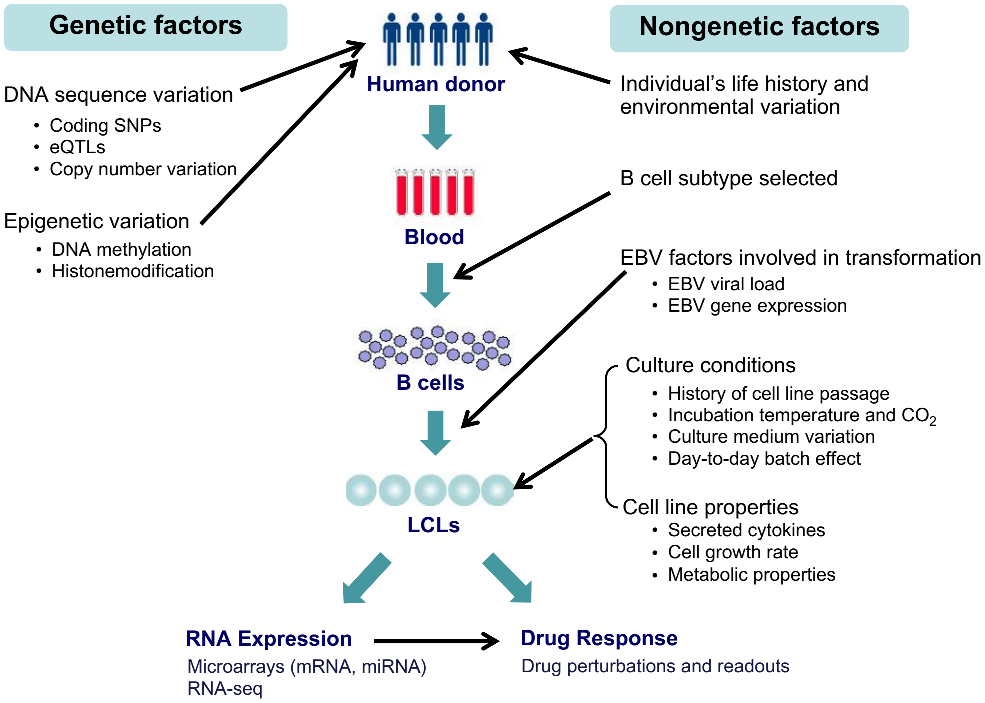

Genetic factors
DNA sequence variation
• Coding SNPs
• eQTLs
• Copy number variation
Epigenetic variation
• DNA methylation
• Histonemodification
Human donor
Blood
B cells
LCLs
Nongenetic factors
Individual's life history and environmental variation
B cell subtype selected
EBV factors involved in transformation
• EBV viral load
• EBV gene expression
Culture conditions
• History of cell line passage
• Incubation temperature and CO2
• Culture medium variation
• Day-to-day batch effect
Cell line properties
• Secreted cytokines
• Cell growth rate
• Metabolic properties
RNA Expression
Microarrays (mRNA, miRNA)
RNA-seq
Drug Response
Drug perturbations and readouts

FIGURE 6.3 **Genetic and nongenetic confounding factors that affect LCL *in vitro* assays.** *Source: Adapted from Choy et al.* [96].

The LCL-based cell model has many advantages over clinical settings, and LCLs are an unlimited source of cellular and genomic materials. For example, LCLs can be easily generated from small amounts of peripheral blood and grown in well-controlled culture conditions, which allows genetic components to be tested without *in vivo* confounders, including comorbidities, dosage, timing, and diet, as well as population stratification or admixture. Moreover, LCLs allow phenotypic measurements (e.g., sensitivities to radiation or chemotherapy) that would be considered unsafe in healthy volunteers or patients. Table 6.2 summarizes some of the advantages and limitations of pharmacogenomic studies in clinical settings and cell-based models [97].

However, the LCL-based assay model has limitations in pharmacogenomic studies. A major limitation is that LCLs represent only one specific lymphoid tissue type. Although some reports have shown that LCLs can be used to discover genetic components or target genes of nonlymphoid neurological diseases such as schizophrenia and autism, LCLs' single lymphoid tissue specificity has hampered their use in pharmacokinetic studies that should be carried out on the human body or in different cell types and organs. This limitation may be overcome in certain cases if LCLs are reprogrammed to various cell and tissue types by iPSC technology.

Other limitations include EBV transformation effects and nongenetic confounding factors, as discussed in the previous section. Such biological noise and *in vitro* artifacts may affect phenotype measurements and subsequent analyses. Taking these limitations into account, initial pharmacogenomic findings using LCLs should be replicated in relevant tissues with functional validation studies, and tested for recapitulation in clinical populations.

6.5 FUTURE PERSPECTIVES

Modern health sciences need an unprecedented number of samples and their associated health data from general populations or disease patients. Well-organized biobanks and information derived from omics and health data have become necessary for the discovery and validation of biomarkers. In the era of big data, personal genetic data, health information, and environmental exposure data will be vigorously combined to accelerate the realization of predictive, personalized, preventive, and participatory medicine [98]. Moreover, statistical power will be crucial in genetic studies of common complex diseases as well as rare diseases and rare drug responses. Thus, harmonization and data sharing will be necessary to achieve the statistical power sufficient for biobank research, including population genetics and pharmacogenomics.

Until now, most cases of LCL applications in pharmacogenomic studies have utilized growth inhibition profiles of drug-treated LCLs to assess cytotoxicity of drugs of interest or environmental stresses such as radiation. Although

TABLE 6.2 Advantages and Limitations of Human and Cell Models

In Vivo Clinical Samples	*In Vitro* Cell Models
Advantages	
• Most relevant study system • Pharmacokinetics as well as pharmacodynamic testing possible • Ability to analyze effects of interactions of multiple tissues • Allows investigation of environmental factors and drug-drug interactions	• Ease of experimental manipulation • Controlled testing system free of *in vivo* confounders • Ability to test drugs with narrow therapeutic index • Genetic and expression data publicly available for hundreds of established cell lines • Cost-effective • Unlimited resources of cellular and genomic materials • Can be used for functional validation of genetic variants
Limitations	
• Difficult to collect large cohorts of patients for discovery and replication • Can be difficult to parse effects of environment from genetic effects • Ethical issues with studying drugs with narrow therapeutic indices • Not all cohorts have detailed pharmacological data, especially retrospective cohorts • Expensive and time-consuming • Require large numbers of patients and infrastructure to obtain reliable clinical phenotype data • Uncontrolled confounders (comorbidities, dosage, timing, diet, population stratification)	• Phenotype may not recapitulate *in vivo* system • Subject to *in vitro* confounders that may affect phenotypes • Cell lines can be generated only from a limited number of tissues • Cell may not express important metabolic enzymes relevant to pharmacology (e.g., cytochrome P450 family) • Establishing cell lines from tumor may alter tumor characteristics

Source: Adapted from [97].

the cytotoxicity assay is technically easy to use in studies of interindividual drug response variation in LCLs, cellular responses to drugs or environmental stresses in LCLs can be assessed by measuring other cellular phenotypes, such as intracellular ATP, Ca2+ levels or Ca2+ efflux, intracellular oxygen levels, or immunophenotypes. Developing such *in vitro* assay models for various cellular phenotypes will expand LCL applications to pharmacogenomics as well as epidemiological genomics.

The exploitation of LCLs' antigen-presenting capacity may provide *in vitro* human therapeutic antibodies against autoimmune and inflammatory diseases, as well as cancer, if the LCLs are generated using resting B lymphocytes from cancer patients. The CD40-CD154 culture system has been used to activate human B lymphocytes, resulting in the generation and clonal expansion of human B-cell lines [99]. Generally, therapeutic intravenous immunoglobulins (IVIgs) prepared from blood donors have been used to re-establish the immune system's homeostasis [99,100]. Moreover, autologous LCLs can be genetically modified to express antigens or mutated genes by transfection methods, including the virus-free episomal method.

Autologous LCLs are a good source of antigen-presenting cells for immunotherapy against cancers and immunological diseases. For example, genetically modified autologous LCLs expressing mutated Ras (muRas) successfully induced muRas-specific cytotoxic T cells in patients with muRas-related pancreatic carcinoma. On the other hand, it is possible that LCLs are generated from patients recovered from a specific disease and then clonally selected for specific autoantigens, providing a new approach to serotherapy using autologous LCLs. Indeed, the use of LCLs as a source of serotherapy has been successfully applied to isolate broadly neutralizing antibodies against SARS and H5N1 viruses [101]. These antibody-producing LCL clones open up novel applications in the field of infectious disease as well as chronic inflammatory and autoimmune disease and cancer.

LCL-derived iPSCs will provide *in vitro* tissue models to aid discovery of molecular targets of drugs and environmental stimuli that may function specifically in different target tissue or cell types (Figure 6.4). Moreover, patient LCL-derived iPSCs not only will facilitate the screening of novel drugs that can be assayed for drug effects on patient-based cell models reflecting ethnic diversity; they may also facilitate development of *in vitro* assay models to investigate cell type–specific adverse drug effects in populations. It is possible that LCL-derived iPSCs will solve the problem of representing LCLs' single lymphoid tissue specificity, which has thus far limited wide use of LCLs. They will extend the utility and application of LCLs to nonlymphoid tissues for different diseases, including brain tissue for neurodegenerative diseases, liver tissue and pancreas tissue for diabetes, and skin tissue, and they may revolutionize pharmacogenomic assay tools for personalized medicine.

Pathogens generally represent tissue specificity or cell tropism for productive life cycles. Although LCLs are derived from lymphocytes, a recent study showed the possibility that they can be used as a system for discovery of pathogen susceptibility biomarkers. If biological activities or infectivity of microbial pathogens can be fully or partially monitored in LCLs, infectious disease susceptibility may be easily investigated using them [57]. It is thus important to

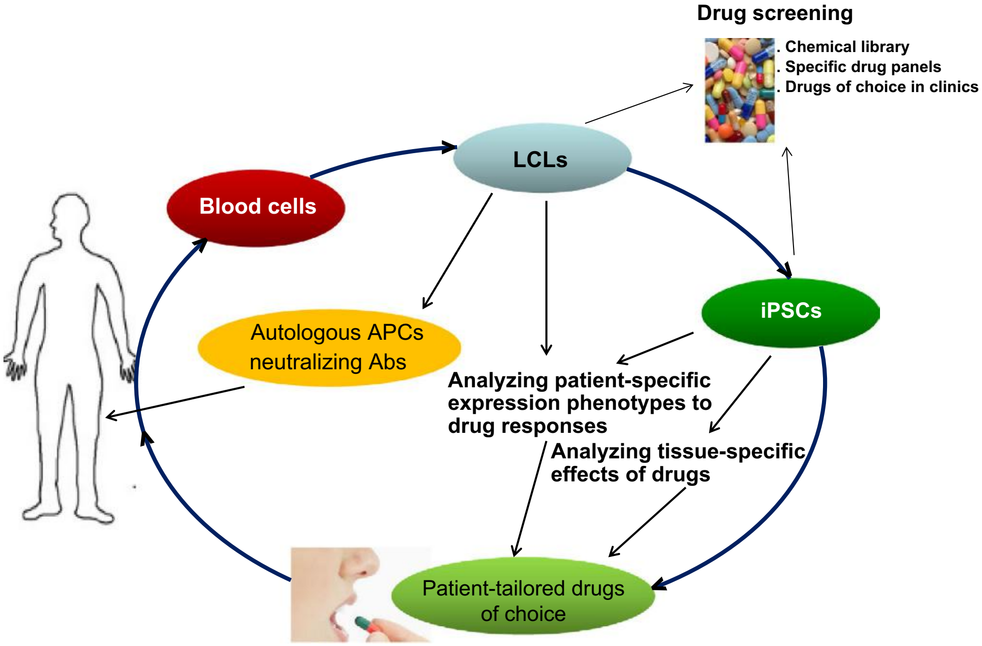

Drug screening
. Chemical library
. Specific drug panels
. Drugs of choice in clinics
LCLs
Blood cells
iPSCs
Autologous APCs
neutralizing Abs
Analyzing patient-specific expression phenotypes to drug responses
Analyzing tissue-specific effects of drugs
Patient-tailored drugs of choice

FIGURE 6.4 Application of LCL-derived iPSCs in personalized medicine.

develop an appropriate *in vitro* cell-based assay system to monitor host–pathogen interaction for particular pathogens in LCLs.

In general, it is easy to establish high-throughput assay systems using a single cell line. For example, a single cell type is first genetically modified to generate the reporter cell line and then grown in microtiter plates for high-throughput screening of multiple drugs. In contrast, with LCLs it is difficult to establish microtiter plates or cell arrays for high-throughput assays because multiple LCL strains are not easy to manipulate in culture at the same time. A novel LCL-based cell array therefore needs to be developed to promote high-throughput manipulation of LCLs to facilitate extended LCL application.

In the personalized medicine era, LCL utility will be extended to various fields, including a cell-based model system in discovery science and therapeutic tools in clinical settings. An international and regional consortium for LCL research may accelerate studies and applications using LCLs as well as LCL harmonization and data sharing.

ACKNOWLEDGMENTS

I would like to thank Dr. David Gurwitz at NLGIP for recommending me to write this book chapter and for his helpful advice to my LCL research. I am also grateful to Dr. Bok-Ghee Han and Dr. Sang-Ick Park for their support of the LCL research in KNIH. This work was supported by intramural grants (2013-NG74001-00, 2012-NG62002-00) from the Korea National Institute of Health and the Korea Centers for Disease Control and Prevention.

REFERENCES

[1] Aggarwal D, Freake HC, Soliman GA, Dutta A, Fernandez ML. Validation of using gene expression in mononuclear cells as a marker for hepatic cholesterol metabolism. Lipids Health Dis 2006;5:22.21–4.

[2] Aman P, Ehlin-Henriksson B, Klein G. Epstein-Barr virus susceptibility of normal human B lymphocyte populations. J Exp Med 1984;159:208–20.

[3] Forsyth NR, Morales CP, Damle S, Boman B, Wright WE, Kopelovich L, et al. Spontaneous immortalization of clinically normal colon-derived fibroblasts from a familial adenomatous polyposis patient. Neoplasia 2004;6:258–65.

[4] Wiesner M, Zentz C, Mayr C, Wimmer R, Hammerschmidt W, Zeidler R, et al. Conditional immortalization of human B cells by CD40 ligation. PLoS ONE 2008;3:e1464.

[5] Amoli MM, Carthy D, Platt H, Ollier WE. EBV Immortalization of human B lymphocytes separated from small volumes of cryo-preserved whole blood. Int J Epidemiol 2008;37(Suppl. 1):i41–5.

[6] Anderson MA, Gusella JF. Use of cyclosporin A in establishing Epstein-Barr virus-transformed human lymphoblastoid cell lines. In Vitro 1984;20(11):856–8.

[7] Baik SY, Yun HS, Lee HJ, Lee MH, Jung SE, Kim JW, et al. Identification of stathmin 1 expression induced by Epstein-Barr virus in human B lymphocytes. Cell Prolif 2007;40(2):268–81.

[8] Takimoto T, Tanaka S, Masuda K, Umeda R. Spontaneous lymphoblastoid transformation of cells obtained from human tonsils. ORL J Otorhinolaryngol Relat Spec 1989;51(2):77–82.

[9] Czerska EM, Elson EC, Davis CC, Swicord ML, Czerski P. Effects of continuous and pulsed 2450-MHz radiation on spontaneous lymphoblastoid transformation of human lymphocytes *in vitro*. Bioelectromagnetics 1992;13(4):247–59.

[10] Kubuschok B, Schmits R, Hartmann F, Cochlovius C, Breit R, König J, et al. Use of spontaneous Epstein-Barr virus-lymphoblastoid cell lines genetically modified to express tumor antigen as cancer vaccines: mutated p21 ras oncogene in pancreatic carcinoma as a model. Hum Gene Ther 2002;13(7):815–27.

[11] Epstein MA, Achong BG, Barr YM. Virus particles in cultured lymphoblasts from Burkitt's lymphoma. Lancet 1964;28:702–3.

[12] Gerber P, Whang-Peng J, Monroe JH. Transformation and chromosome changes induced by Epstein-Barr virus in normal human leukocyte cultures. Proc Natl Acad Sci USA 1969;63:740–7.

[13] Pope JH, Horne MK, Scott W. Transformation of foetal human keukocytes *in vitro* by filtrates of a human leukaemic cell line containing herpes-like virus. Int J Cancer 1968;3(6):857–66.

[14] Nemerow GR, Wolfert R, McNaughton ME, Cooper NR. Identification and characterization of the Epstein-Barr virus receptor on human B-lymphocytes and its relation to the C3d complement receptor (CR2). J Virol 1985;55:347.

[15] Klein G, Giovanella B, Westman A, Stehlin JS, Mumford D. An EBV-genome-negative cell line established from an American Burkitt lymphoma; receptor characteristics. EBV infectibility and permanent conversion into EBV-positive sublines by *in vitro* infection. Intervirology 1975;5:319.

[16] Klein G, Lindahl T, Jondal M, Leibold W, Menézes J, Nilsson K, et al. Continuous lymphoid cell lines with characteristics of B cells (bone-marrow-derived), lacking the Epstein-Barr virus genome and derived from three human lymphomas. Proc Natl Acad Sci USA 1974;71:3283.

[17] Pattengale PK, Smith RW, Gerber P. Selective transformation of B lymphocytes by E.B. virus. Lancet 1973;2(7820):93–4.

[18] Henderson E, Miller G, Robinson J, Heston L. Efficiency of transformation of lymphocytes by Epstein-Barr virus. Virology 1977;76:152.

[19] Nilsson K, Klein G. Phenotypic and cytogenetic characteristics of human B-lymphoid cell lines and their relevance for the etiology of Burkitt's lymphoma. Adv Cancer Res 1982;37:319.

[20] Bass H, Walters R, Darke C. Provision of Epstein-Barr virus-transformed B-cell lines in a routine tissue typing laboratory: practicalities and applications. Eur J Immunogenet 2004;31(2):87–92.

[21] Neitzel H. A routine method for establishment of permanent growing lymphoblastoid cell lines. Hum Genet 1986;73:320–6.

[22] Tosato G, Cohen JI. Generation of Epstein-Barr virus (EBV)-immortalzied B cell lines. Curr Protoc Immunol 2007:7.22.1–7.22.4.

[23] Benyesh-Melnick M, Fernbach DJ, Lewis RT. Studies on human Leukemia. I. Spontaneous lymphoblastoid transformation of fibroblastic bone marrow cultures derived from leukemic and nonleukemic children. JNCI J Natl Cancer Inst 1963;31(6):1311–25.

[24] Miller G, Lipman M. Release of infectious Epstein-Barr virus by transformed marmoset leukocytes. Proc Natl Acad Sci USA 1973;70:190–4.

[25] Louie LG, King MC. A novel approach to establishing permanent lymphoblastoid cell lines: Epstein-Barr virus transformation of cryopreserved lymphocytes. Am J Hum Genet 1991;48:637–8.

[26] Pressman S, Jerome I, Rotter JI. Epstein-Barr virus transformation of cryopreserved lymphocytes: prolonged experience with technique. Am J Hum Genet 1991;49:467.

[27] Hui-Yuen J, McAllister S, Koganti S, Hill E, Bhaduri-McIntosh S. Establishment of Epstein-Barr virus growth-transformed lymphoblastoid cell lines. J Vis Exp 2011;(57):3321. pii

[28] Bird AG, McLachlan SM, Britton S. Cyclosporin A promotes spontaneous outgrowth *in vitro* of Epstein-Barr virus-induced B-cell lines. Nature 1981;289:300–1.

[29] Matsuda S, Koyasu S. Mechanisms of action of cyclosporine. Immunopharmacology 2000;47:119–25.

[30] Megyola C, Ye J, Bhaduri-McIntosh S. Identification of a sub-population of B cells that proliferates after infection with Epstein-Barr virus. Virol J 2011;8:84.

[31] Londin ER, Keller MA, D'Andrea MR, Delgrosso K, Ertel A, Surrey S, et al. Whole-exome sequencing of DNA from peripheral blood mononuclear cells (PBMC) and EBV-transformed lymphocytes from the same donor. BMC Genomics 2011;12:464.

[32] Beck JC, Beiswanger CM, John EM, Satariano E, West D. Successful transformation of cryopreserved lymphocytes: a resource for epidemiological studies. Cancer Epidemiol Biomarkers Prev 2001;10:551–4.

[33] Bernacki SH, Stankovic AK, Williams LO, Beck JC, Herndon JE, Snow-Bailey K, et al. Establishment of stably EBV-transformed cell lines from residual clinical blood samples for use in performance evaluation and quality assurance in molecular genetic testing. J Mol Diagn 2003;5:227–30.

[34] Chatterjee R. Cell biology. Cases of mistaken identity. Science 2007;315:928–31.

[35] MacLeod RA, Dirks WG, Matsuo Y, Kaufmann M, Milch H, Drexler HG. Widespread intraspecies cross-contamination of human tumor cell lines arising at source. Int J Cancer 1999;83:555–63.

[36] Capes-Davis A, Theodosopoulos G, Atkin I, Drexler HG, Kohara A, MacLeod RA, et al. Check your cultures! A list of cross-contaminated or misidentified cell lines. Int J Cancer 2010;127(1):1–8.

[37] Freshney RI. Database of misidentified cell lines. Int J Cancer 2010;126:302.

[38] Drexler HG, Dirks WG, Matsuo Y, MacLeod RA. False leukemia–lymphoma cell lines: an update on over 500 cell lines. Leukemia 2003;17:416–26.

[39] Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, Piouffre L, et al. A human genome diversity cell line panel. Science 2002;296(5566):261–2.

[40] Rosenberg NA. Standardized subsets of the HGDP-CEPH Human Genome Diversity Cell Line Panel, accounting for atypical and duplicated samples and pairs of close relatives. Ann Hum Genet 2006;70:841–7.

[41] Zhang W, Ratain MJ, Dolan ME. The HapMap resource is providing new insights into ourselves and its application to pharmacogenomics. Bioinform Biol Insights 2008;2:15–23.

[42] Nam HY, Shim SM, Han BG, Jeon JP. Human lymphoblastoid cell lines: a goldmine for the biobankomics era. Pharmacogenomics 2011;12(6):907–17.

[43] Herbeck JT, Gottlieb GS, Wong K, Detels R, Phair JP, Rinaldo CR, et al. Fidelity of SNP array genotyping using Epstein Barr virus-transformed B-lymphocyte cell lines: implications for genome-wide association studies. PLoS ONE 2009;4(9):e6915.

[44] Detels R, Jacobson L, Margolick J, Martinez-Maza O, Muñoz A, Phair J, et al. The multicenter AIDS Cohort Study, 1983 to …. Public Health 2012;126(3):196–8.

[45] Fu SM, Hurley JN. Human cell lines containing Epstein-Barr virus but distinct from the common B cell lymphoblastoid lines. Proc Natl Acad Sci USA 1979;76:6637–40.

[46] Kilger E, Kieser A, Baumann M, Hammerschmidt W. Epstein-Barr virus-mediated B-cell proliferation is dependent upon latent

membrane protein 1, which stimulates an activated CD40 receptor. EMBO J 1998;17:1700–9.

[47] Hur DY, Lee MH, Kim JW, Kim JH, Shin YK, Rho JK, et al. CD19 signalling improves the Epstein-Barr virus-induced immortalization of human B cell. Cell Prolif 2005;38(1):35–45.

[48] Oh HM, Oh JM, Choi SC, Kim SW, Han WC, Kim TH, et al. An efficient method for the rapid establishment of Epstein-Barr virus immortalization of human B lymphocytes. Cell Prolif 2003;36(4):191–7.

[49] Carter RH, Fearon DT. CD19: lowering the threshold for antigen receptor stimulation of B lymphocytes. Science 1992;256:05–107.

[50] Sugimoto M, Tahara H, Ide T, Furuichi Y. Steps involved in immortalization and tumorigenesis in human B-lymphoblastoid cell lines transformed by Epstein-Barr virus. Cancer Res 2004;64:3361–4.

[51] Jeon JP, Nam HY, Shim SM, Han BG. Sustained viral activity of Epstein-Barr virus contributes to cellular immortalization of lymphoblastoid cell lines. Mol Cells 2007;27:143–8.

[52] Jeon JP, Shim SM, Nam HY, Baik SY, Kim JW, Han BG. Copy number increase of 1p36.33 and mitochondrial genome amplification in Epstein-Barr virus-transformed lymphoblastoid cell lines. Cancer Genet Cytogenet 2007;173(2):122–30.

[53] Lee JE, Nam HY, Shim SM, Bae GR, Han BG, Jeon JP. Expression phenotype changes of EBV-transformed lymphoblastoid cell lines during long-term subculture and its clinical significance. Cell Prolif 2010;43:378–84.

[54] Lee JE, Hong EJ, Nam HY, Hwang M, Kim JH, Han BG, et al. Molecular signatures in response to isoliquiritigenin in lymphoblastoid cell lines. Biochem Biophys Res Commun 2012;427:392–7.

[55] Caliskan M, Cusanovich DA, Ober C, Gilad Y. The effects of EBV transformation on gene expression levels and methylation profiles. Hum Mol Genet 2011;20(8):1643–52.

[56] Oh JH, Kim YJ, Moon S, Nam HY, Jeon JP, Lee JH, et al. Genotype instability during long-term subculture of lymphoblastoid cell lines. J Hum Genet 2013;58(1):16–20.

[57] Loeuillet C, Deutsch S, Ciuffi A, Robyr D, Taffé P, Munoz M, et al. *In vitro* whole-genome analysis identifies a susceptibility locus for HIV-1. PLoS Biol 2008;6(2):e32.

[58] Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature 2006;444(7118):444–54.

[59] Simon-Sanchez J, Scholz S, Fung HC, Matarin M, Hernandez D, Gibbs JR, et al. Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. Hum Mol Genet 2007;16:1–14.

[60] Jeon JP, Shim SM, Nam HY, Ryu GM, Hong EJ, Kim HL, et al. Copy number variation at leptin receptor gene locus associated with metabolic traits and the risk of type 2 diabetes mellitus. BMC Genomics 2010;12:11–426.

[61] Nickles D, Madireddy L, Yang S, Khankhanian P, Lincoln S, Hauser SL, et al. In depth comparison of an individual's DNA and its lymphoblastoid cell line using whole genome sequencing. BMC Genomics 2012;14:13–477.

[62] Hu VW, Sarachana T, Kim KS, Nguyen A, Kulkarni S, Steinberg ME, et al. Gene expression profiling differentiates autism case-controls and phenotypic variants of autism spectrum disorders: evidence for circadian rhythm dysfunction in severe autism. Autism Res 2009;2(2):78–97.

[63] Jen KY, Cheung VG. Transcriptional response of lymphoblastoid cells to ionizing radiation. Genome Res 2003;13:2092–100.

[64] Sawada K, Noda K, Nakajima H, Shimbara N, Furuichi Y, Sugimot M. Differential cytotoxicity of anticancer agents in pre- and post-immortal lymphoblastoid cell lines. Biol Pharm Bull 2005;28:1202–7.

[65] Gimelbrant A, Hutchinson JN, Thompson BR, Chess A. Widespread monoallelic expression on human autosomes. Science 2007;318:1136–40.

[66] Plagnol V, Uz E, Wallace C, Stevens H, Clayton D, Ozcelik T, et al. Extreme clonality in lymphoblastoid cell lines with implications for allele specific expression analyses. PLoS ONE 2008;3(8):e2966.

[67] Wu L, Candille SI, Choi Y, Xie D, Jiang L, Li-Pook-Than J, et al. Variation and genetic control of protein abundance in humans. Nature 2013;499:79–82.

[68] Zhijian C, Xiaoxue L, Wei Z, Yezhen L, Jianlin L, Deqiang L, et al. Studying the protein expression in human B lymphoblastoid cells exposed to 1.8-GHz (GSM) radiofrequency radiation (RFR) with protein microarray. Biochem Biophys Res Commun 2013;433:36–9.

[69] Kazuno AA, Ohtawa K, Otsuki K, Usui M, Sugawara H, Okazaki Y, et al. Proteomic analysis of lymphoblastoid cells derived from monozygotic twins discordant for bipolar disorder: a preliminary study. PLoS ONE 2013;8(2):e53855.

[70] Docherty SJ, Davis OS, Haworth CM, Plomin R, D'Souza U, Mill J. A genetic association study of DNA methylation levels in the DRD4 gene region finds associations with nearby SNPs. Behav Brain Funct 2012;8:31.

[71] Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol 2011;12(1):R10.

[72] Toung JM, Morley M, Li M, Cheung VG. RNA-sequence analysis of human B-cells. Genome Res 2011;21:991–8.

[73] Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. Nature 2010;464(7289):773–7.

[74] Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature 2010;464(7289):768–72.

[75] Skalsky RL, Corcoran DL, Gottwein E, Frank CL, Kang D, Hafner M, et al. The viral and cellular MicroRNA targetome in lymphoblastoid cell lines. PLoS Pathog 2012;8(1):e1002484.

[76] Njiaju UO, Gamazon ER, Gorsic LK, Delaney SM, Wheeler HE, Im HK, et al. Whole-genome studies identify solute carrier transporters in cellular susceptibility to paclitaxel. Pharmacogenet Genomics 2012;22:498–507.

[77] Huang RS, Duan S, Kistner EO, Hartford CM, Dolan ME. Genetic variants associated with carboplatin-induced cytotoxicity in cell lines derived from Africans. Mol Cancer Ther 2008;7:3038–46.

[78] Li L, Fridley BL, Kalari K, Jenkins G, Batzler A, Weinshilboum RM, et al. Gemcitabine and arabinosylcytosin pharmacogenomics: genome-wide association and drug response biomarkers. PLoS ONE 2009;4(11):e7765.

[79] Hartford CM, Duan S, Delaney SM, Mi S, Kistner EO, Lamba JK, et al. Population-specific genetic variants important in susceptibility to cytarabine arabinoside cytotoxicity. Blood 2009;113:2145–53.

[80] Weng L, Ziliak D, Im HK, Gamazon ER, Philips S, Nguyen AT, et al. The consortium on breast cancer pharmacogenomics (COBRA). Genome-wide discovery of genetic variants affecting

tamoxifen sensitivity and their clinical and functional validation. Ann Oncol 2013;24(7):1867–73.

[81] Niu N, Qin Y, Fridley BL, Hou J, Kalari KR, Zhu M, et al. Radiation pharmacogenomics: a genome-wide association approach to identify radiation response biomarkers using human lymphoblastoid cell lines. Genome Res 2010;20:1482–92.

[82] Morag A, Pasmanik-Chor M, Oron-Karnl V, Rehavi M, Stingl JC, Gurwitz D. Genome-wide expression profiling of human lymphoblastoid cell lines identifies CHL1 as a putative SSRI antidepressant response biomarker. Pharmacogenomics 2011;12:171–84.

[83] Fry RC, Sensson JP, Valiathan C, Wang E, Hogan BJ, Bhattacharya S, et al. Genomic predictors of interindividual differences in response to DNA damaging agents. Genes Dev 2008;22:2621–6.

[84] Huang RS, Johnatty SE, Gamazon ER, Im HK, Ziliak D, Duan S, et al. Platinum sensitivity-related germline polymorphism discovered via a cell-based approach and analysis of its association with outcome in ovarian cancer patients. Clin Cancer Res 2011;17(16):5490–500.

[85] Medina MW, Gao F, Ruan W, Rotter JI, Krauss RM. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. Circulation 2008;118(4):355–62.

[86] Kuppers R. B cells under influence: transformation of B cells by Epstein-Barr virus. Nat Rev Immunol 2003;3(10):801–12.

[87] Hurley JN, Fu SM, Kunkel HG, McKenna G, Scharff MD. Lymphoblastoid cell lines from patients with chronic lymphocytic leukemia: identification of tumor origin by idiotypic analysis. Proc Natl Acad Sci USA 1978;75:5706–10.

[88] Steinitz M, Klein G, Koskimies S, Makel O. EB virus-induced B lymphocyte cell lines producing specific antibody. Nature 1977;269:420–2.

[89] Traggiai E, Becker S, Subbarao K, Kolesnikova L, Uematsu Y, Gismondo MR, et al. An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat Med 2004;10:871–5.

[90] Neumann F, Kaddu-Mulindwa D, Widmann T, Preuss KD, Held G, Zwick C, et al. EBV-transformed lymphoblastoid cell lines as vaccines against cancer testis antigen-positive tumors. Cancer Immunol Immunother 2013;62:1211–22.

[91] Sun Q, Burton RL, Pollok LE, Emanuel DJ, Lucas KG. CD(+) Epstein-Barr virus-specific cytotoxic T-lymphocytes from human umbilical cord blood. Cell Immunol 1999;195(2):81–8.

[92] Okugawa K, Itoh T, Kawashima I, et al. Recognition of Epstein-Barr virus-associated gastric T lymphocytes induced *in vitro* with autologous lymphoblastoid cell line and LMP2-derived, HLA-A24-restriced 9-mer peptide. Oncol Rep 2004; 12(4):725–31.

[93] Choi SM, Liu H, Chaudhari P, Kim Y, Cheng L, Feng J, et al. Reprogramming of EBV-immortalized B-lymphocyte cell lines into induced pluripotent stem cells. Blood 2011;118(7):1801–5.

[94] Rajesh D, Dickerson SJ, Yu J, Brown ME, Thomson JA, Seay NJ. Human lymphoblastoid B-cell lines reprogrammed to EBV-free induced pluripotent stem cells. Blood 2011;118:1797–800.

[95] Ye Z, Liu CF, Jang YY. Hematopoietic cells as sources for patient-specific iPSCs and disease modeling. Cell Cycle 2011;10:2840–4.

[96] Choy E, Yelensky R, Bonakdar S, Plenge RM, Saxena R, De Jager PL, et al. Genetic analysis of human traits *in vitro*: drug response and gene expression in lymphoblastoid cell lines. PLoS Genet 2008;4:e1000287.

[97] Welsh M, Mangravite L, Medina MW, Tantisira K, Zhang W, Huang RS, et al. Pharmacogenomics discovery using cell-based models. Pharmacol Rev 2009;61:413–29.

[98] Hood L, Friend SH. Predictive, personalized, preventive, participatory (P4) cancer medicine. Nat Rev Clin Oncol 2011;8(3):184–7.

[99] Neron S, Roy A, Dumont N. Large-scale in vitro expansion of polyclonal human switched-memory B lymphocytes. PLoS ONE 2012;7(12):e51946.

[100] Durandy A, Kaveri SV, Kuijpers TW, Basta M, Miescher S, Ravetch JV, et al. Review Intravenous immunoglobulins–understanding properties and mechanisms. Clin Exp Immunol 2009;158(Suppl 1):2–13.

[101] Lanzavecchia A, Corti D, Sallusto F. Human monoclonal antibodies by immortalization of memory B cells. Curr Opin Biotechnol 2007;18(6):523–8.

Chapter 7

Liver Expression Quantitative Trait Loci (eQTL) and Related Approaches in Pharmacogenomic Studies

Dylan M. Glubb,[a] Amy S. Etheridge,[b] Eric Seiser,[b] and Federico Innocenti[b,c,d]

[a]*Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Australia,* [b]*Eshelman School of Pharmacy, Institute for Pharmacogenomics and Individualized Therapy, University of North Carolina at Chapel Hill, USA,* [c]*Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, USA,* [d]*School of Medicine, University of North Carolina at Chapel Hill, USA*

Chapter Outline

7.1 Overview **111**
7.2 Liver eQTL Studies **112**
  7.2.1 Liver eQTL in Pharmacogenetic and Clinical Genome-Wide Association Studies 112
  7.2.2 Liver eQTL Reproducibility 115
  7.2.3 Liver eQTL Tissue Specificity 116
  7.2.4 Correlation of Liver eQTLs and CYP Enzyme Activity 116
  7.2.5 Web-Based Liver eQTL Data Resources 116
7.3 Future Directions **117**
  7.3.1 Next-Generation Sequencing 117
  7.3.2 Identification of sQTL 118
  7.3.3 Identification of miRNA eQTL 118
  7.3.4 Identification of Functional eSNPs 118
  7.3.5 Epistatic eQTL Interactions 119
  7.3.6 Copy Number Variants 119
  7.3.7 Liver Protein QTL 120
7.4 Conclusion **120**
Acknowledgments **120**
Glossary **121**
Acronyms and Abbreviations **121**
References **121**

7.1 OVERVIEW

Pharmacogenomics is the study of heritable genetic variation in drug response, and aims to elucidate how genetic variants influence drug efficacy and/or toxicity. Human germline genetic variation is well known to affect drug disposition and clinical response to pharmacotherapy via genes related to both therapeutic targets and absorption, distribution, metabolism, and excretion (ADME) [1]. The Pharmacogenomics Knowledgebase (PharmGKB; www.pharmgkb.org) curates pharmacogenetic findings from the literature and, as of September 2013, lists 453 associations between germline genetic variants and drug-related phenotypes. Only 13% of these associations can be explained by genetic variation due to protein alterations as a result of changes in the amino acid sequence. However, genetic variation can impact ADME gene expression and clinical outcomes on numerous levels, including gene transcription, alternative splicing (AS), transcript stability, and rate of translation, as well as protein expression, function, and degradation [2,3]. As a result, mRNA transcript levels provide clearly defined and measurable traits that may act as intermediates for drug-related phenotypes.

Genetic variants that associate with gene expression are known as expression quantitative trait loci (eQTL) and can be further characterized as *cis* or *trans*. *cis*-eQTL are located near the expressed gene, usually within 1 Mb, while *trans*-eQTL are located further away, sometimes on a different

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00007-4
© 2014 Elsevier Inc. All rights reserved.

chromosome than the expressed gene. eQTL can refer to any type of genetic variation, but most often indicate that a single-nucleotide polymorphism (SNP) is associated with the transcript level of a gene. The interrogation of SNP genotypes and transcript levels has generally been performed using oligonucleotide microarray platforms, enabling the identification of eQTL through genome-wide analyses [4]. This linking of gene variation and expression may clarify the pharmacogenetic significance of genes of interest [5] and thus eQTL have significant potential utility for the study of pharmacogenomics.

Lymphoblastoid cell lines (LCLs) have been the predominant model for eQTL analyses in humans [6]. These cell lines are derived from B-lymphocytes transformed by the Epstein-Barr virus (EBV). Although this transformation provides an immortalized cell line that can be directly examined in pharmacogenomic studies [7–9], nongenetic factors—including the life history of the human donor, the B-lymphocyte subtype selected, the EBV titers used for transformation, the culture conditions under which the cells are grown, and the intrinsic characteristics of lymphoblastoid cells—can affect gene expression and limit the utility of LCLs for these studies [10]. Nevertheless, several studies have incorporated LCL eQTL findings in clinical pharmacogenomic studies (reviewed in Wheeler et al. [11]).

In the context of clinical pharmacogenomics, the liver may be the most relevant tissue in which to perform eQTL analyses. It is the predominant organ in drug elimination, with 75% of the 200 most widely prescribed drugs being eliminated from the body through liver metabolism or biliary excretion [12]. Furthermore, a significant number of ADME genes are strongly expressed in the liver. Schroder et al. [13] defined a list of 682 ADME genes from various pharmacogenomic resources. The expression of 61 of these genes has been predicted to be liver-specific according to a study of gene expression in human tissues [14], indicating that eQTL of these genes may not be detected in other tissues. In addition, eQTL may have effects at the molecular level that depend on the cell or organ in question. Therefore, knowledge of liver eQTL could provide a foundation for clinical pharmacogenomic research.

7.2 LIVER eQTL STUDIES

We are presently in the early stages of liver eQTL research, and only five genome-wide eQTL studies have been reported that describe the analysis of primary human liver tissue (Table 7.1). However, these studies have led to the identification of several thousand eQTL, after correction for multiple testing, and many of the eQTL single-nucleotide polymorphisms (eSNPs) associate with the expression of ADME genes.

7.2.1 Liver eQTL in Pharmacogenetic and Clinical Genome-Wide Association Studies

ADME genes with liver eQTL, which pass multiple correction thresholds in one of the five genome-wide association studies (GWAS), include *COMT*, *CYP2D6*, *CYP2E1*, *CYP3A4*, *CYP3A5*, *GSST1*, *MTHFR*, *UGT1A1*, and *VKORC1* (Table 7.2). These genes belong to a set of 50 "very important pharmacogenes" (VIPs) that are known to effect drug response phenotypes and have been identified by the PharmGKB resource (http://www.pharmgkb.org/search/browseVip.action?browseKey=vipGenes) as having particular relevance to pharmacogenetics and pharmacogenomics. In addition, several of the eSNPs identified in these studies are markers of pharmacogenetic associations (Table 7.2). These findings demonstrate that liver eQTL may help explain the association of SNPs with clinical phenotypes and provide a rationale for using liver eQTL data from relevant genes to select candidate SNPs for pharmacogenomic study. For example, Schadt et al., [15] identified an eSNP in *COMT* (Table 7.2). This gene encodes catechol O-methyltransferase (COMT), an enzyme that metabolizes catecholamine neurotransmitters, including dopamine, epinephrine, and norepinephrine; also, a *COMT* SNP has been associated with antidepressant drug response [16]. COMT protein is known to exist in two forms, resulting from the use of alternative promoters, and the enzymatic activity of these two isoforms has been demonstrated to vary as a result of altered protein expression.

The role of *COMT* in treatment-resistant major depressive disorder is thought to be related to decreased norepinephrine and dopamine availability resulting from expression of the high-activity enzyme. Therefore, determining the effect of this liver eQTL on mRNA transcript levels, as an intermediary to protein isoform expression, may offer an explanation for these differences in metabolism and thus a basis for the association of this gene with psychiatric phenotypes. Cytochrome P450 (CYP, P450) 2D6 is another important drug-metabolizing enzyme, with the PharmGKB estimating that it is involved in the metabolism of up to 25% of commonly prescribed therapeutics [17]. Additionally, CYP2D6 is the only noninducible drug-metabolizing CYP; therefore, genetic variation contributes significantly to interindividual variability in the activity of this enzyme. Three eSNPs have been identified in the *CYP2D6* gene as a result of liver eQTL studies (Table 7.2) and each could conceivably be used to test associations with pertinent clinical phenotypes of any of the 46 drugs that, according to PharmGKB-curated information (http://www.pharmgkb.org/gene/PA128#tabview=tab6&subtab=21), are metabolized by the CYP2D6 enzyme.

In contrast to CYP2D6, CYP2E1 makes up a much smaller proportion (<1%) of total hepatic P450 and plays a more minor role in the metabolism of clinical drugs. The anti-tuberculosis drug isoniazid, however, has been shown to be both a substrate for and an inducer of CYP2E1 protein, although it does not increase mRNA levels. Based on these findings, it is thought that protein expression levels may be regulated by increased translational efficiency

TABLE 7.1 Summary of Human Liver eQTL Studies

Category	Schadt et al.	Innocenti et al.	Schroder et al.	Greenawalt et al.	Fu et al.
Samples in final analysis	384	206	149	651	74
Race	All Caucasian	183 European American; 23 African American	All Caucasian	86% of overall cohort Caucasian	All Caucasian
Nature of liver tissue	Postmortem tissue and resections from donor livers	Postmortem tissue and resections from donor livers	Normal tissue resected during surgery for liver cancer	Tissue from morbidly obese patients undergoing gastric bypass	Liver wedge biopsies from morbidly obese patients undergoing gastric bypass
Covariates	Age, sex, medical center	Age, sex, genetic ancestry, effects of array hybridization and oligonucleotide probes, surrogate variables	Age, sex, smoking, alcohol consumption, diagnosis, C-reactive protein, cholestatic liver disease, presurgical medication	Age, sex, medical center	Removed first 50 PCs
Genotyping platform (GEO dataset #)	Affymetrix 500k Illumina 650Y (GSE # n/a)	Illumina 610 Quad (GSE26105)	Illumina HumanHap300 (GSE39036)	Illumina 650Y (GSE # n/a)	Illumina Human Omni1-Quad (GSE # n/a)
Expression platform (GEO dataset #)	Agilent Custom 44k (GSE9588)	Agilent 4×44k (GSE25935)	Illumina Human WG-6 v2.0 (GSE32504)	Agilent Custom 44k (GSE24293)	Illumina Human HT12 (GSE22070)
cis-eQTL #	1350 (Bonferroni-adjusted) 3210 (FDR < 10%)	1787 (BF > 5)	1179 (Bonferroni-adjusted)	7902 (FDR < 10%)	4700 (FDR < 5%)
trans-eQTL #	242 (Bonferroni-adjusted) 491 (FDR < 10%)	353 (BF > 5)	47 (Bonferroni-adjusted)	785 (FDR < 10%)	n/a

and/or protein stabilization. eSNPs have been demonstrated to exert regulatory effects on gene expression post-transcriptionally via mechanisms such as AS and altered 3′-untranslated regions (UTRs). Therefore, a mechanistic role to explain the increased risk of isoniazid-induced liver toxicity associated with *CYP2E1* variants having increased enzyme activity and inducibility could be investigated for the liver eQTL identified by Schadt et al. At the opposite end of the spectrum of human hepatic P450s from *CYP2E1* is the *CYP3A* family, including *CYP3A4* and *CYP3A5*. The proteins encoded by these genes together account for approximately 30% of total human hepatic CYPs and are responsible for the metabolism of >50% of all currently prescribed pharmaceuticals. PharmGKB has curated information for 72 drugs. (http://www.pharmgkb.org/gene/PA130?tabType=tabVip#tabview=tab6&subtab=21) and 67 (http://www.pharmgkb.org/gene/PA131?tabType=tabVip#tabview=tab6&subtab=21) drugs that are substrates for CYP3A4 and CYP3A5, respectively, although significant overlap in specificity exists. Numerous pharmacogenetic associations between *CYP3A4/5* SNPs and drug efficacy or toxicity phenotypes have been identified and, as was described for *CYP2D6*, the *CYP3A4* and 2 *CYP3A5* eSNPs identified from liver eQTL studies could be tested for association with relevant clinical outcomes with any of the known drug substrates.

Phase II enzymes also play a prominent role in pharmacogenetic phenotypes by facilitating detoxification through conjugation reactions that increase solubility and thus aid the elimination of conjugated compounds from the body. Glutathione S-transferase (GST) theta (GSTT1) and uridine diphosphate glucuronosyltransferase (UGT) 1A1 catalyze the conjugation of a number of chemotherapeutics with glutathione and glucuronic acid, respectively, and expression level variation in these genes represents significant potential for altered efficacy and/or toxicity of cancer therapies.

TABLE 7.2 Liver eQTLs of PharmGKB "Very Important Pharmacogenes"

Pharmacogene	eSNP	eQTL Study	PharmGKB Pharmacogenetic SNPs[1] Sharing Perfect LD[2] ($r^2=1$) with Liver eSNP	
			Pharmacogenetic SNP	*Related Drug(s)*
COMT	rs174674	Schadt et al.	–	–
CYP2D6	rs151096	Schadt et al.	–	–
	rs8138080	Schadt et al.	–	–
	rs133337	Innocenti et al.	–	–
CYP2E1	rs4512750	Schadt et al.		
CYP3A4	rs7792939	Greenawalt et al.		–
CYP3A5	rs10242455	Schadt et al.	rs4646457	Tacrolimus
		Schroder et al.	rs776746	Cyclosporine
				Nevirapine
				Paclitaxel
				Rosuvastatin
				Sirolimus
				Sunitinib
				Tacrolimus
	rs1859690	Schadt et al.	rs4646457	Tacrolimus
		Schroder et al.	rs776746	Tacrolimus
				Cyclosporine
				Nevirapine
				Paclitaxel
				Rosuvastatin
				Sirolimus
				Sunitinib
GSTT1	rs4822458	Schadt et al.		
	rs1007888	Schroder et al.		
MTHFR	rs1801131	Innocenti et al.	rs1801131	Methotrexate
				Nitrous oxide
UGT1A1	rs2070959	Schroder et al.	–	–
VKORC1	rs2303222	Innocenti et al.	rs2359612	Warfarin
			rs8050894	Warfarin
	rs10871454	Schroder et al.	rs2359612	Warfarin
			rs8050894	Warfarin
	rs889548	Schroder et al.	–	–
	rs749767	Schroder et al.	–	–
	rs4889606	Schadt et al.	–	–

[1] *From the PharmGKB curated list of level 1–3 SNP-drug clinical annotations (http://www.pharmgkb.org/search/clinicalAnnotationList.action?evidenceStrength=1).*

[2] *LD determined from 1000 Genomes Pilot 1 CEU data using SNAP browser (http://www.broadinstitute.org/mpg/snap/ldsearch.php).*

Note: PharmGKB currently identifies 50 genes of particular importance in pharmacogenomics (http://www.pharmgkb.org/search/browseVip.action?browseKey=vipGenes).

Although more than 130 variants in *GSTT1* are cataloged in the National Center for Biotechnology Information (NCBI) Single Nucleotide Polymorphism Database (dbSNP) [18] as of September 2013, the relevance of most of these SNPs is unknown, as functional roles for only a few have been established through studies reported in the literature and only one pharmacogenetic association has been identified to date. However, a number of associations have been identified between the *GSTT1* null variants (i.e., the partial or complete deletion of the *GSTT1* gene) and adverse drug reactions [17]. Gamazon et al. [19] have shown that SNPs that are in strong linkage disequilibrium (LD) ($r^2 = 0.8$) with copy number variants (CNVs) are more likely to be LCL eQTL. This highlights an area for further investigation of the liver *GSTT1* eSNPs identified by Schadt et al. [15] and Schroder et al. [13].

As opposed to the paucity of data in the literature on genetic variants and pharmacogenetic associations in *GSTT1*, there are 113 known genetic variants in *UGT1A1*, with a number of these known to associate with the development of drug toxicities [17]. In Schroder et al.'s study [13], rs2070959 was identified as a *UGT1A1* eSNP. This variant has not been well studied, and the role it plays in altered UGT1A1 enzymatic phenotypes and possibly the prediction of drug toxicities is yet to be explored. Finally, five liver eQTL were identified in *VKORC1,* a gene with polymorphisms known to account for approximately 25% of interindividual variability in warfarin dose. Two of these eSNPs were in perfect LD ($r^2 = 1$) with PharmGKB pharmacogenetic SNPs, while the remaining three have not been identified as markers of known pharmacogenetic associations. SNPs in *VKORC1* are associated with both high and low warfarin dose phenotypes, and liver eQTL may aid in explaining the role of gene regulation in this variability.

In addition to helping explain the association of SNPs with clinical phenotypes, liver eQTL data can be used to inform GWAS. GWAS have identified many SNPs associated with clinical traits in humans. Typically, there is no mechanistic explanation for the association, but eQTL could help fill this knowledge gap. In the liver eQTL studies conducted to date (Table 7.1), Greenawalt et al. [20] and Schadt et al. [15] identified 121 and 21 eSNPs, respectively, which were significantly associated, after multiple test correction, with a clinical phenotype in a GWAS. Innocenti et al. [21] identified 32 eSNPs that are significant GWAS SNPs or markers for a GWAS SNP through strong LD ($r^2 > 0.8$). In fact, there is a significant enrichment of eSNPs among SNPs associated with clinical phenotypes in GWAS [22]. However, a critical piece of information lacking from these reports is the directionality of the eQTL association (i.e., which allele associates with higher/lower transcript levels), and thus hypotheses about the effect of a specific allele cannot be generated.

For GWAS SNPs that are associated with a clinical phenotype and located in intergenic regions, there is often ambiguity about which gene the SNP is regulating. eQTL data can also be used to aid this determination. Identifying gene expression levels associated with a GWAS SNP provides evidence to suggest that a gene may be a susceptibility or protective factor for the clinical phenotype. Schadt et al. [15] provide examples of intergenic eSNPs that were found to associate with type 1 diabetes, coronary artery disease, and LDL cholesterol levels in GWAS.

7.2.2 Liver eQTL Reproducibility

There are many factors related to liver eQTL studies that could act as confounders and affect eQTL reproducibility. First, these studies used a variety of tissue sources from individuals with different ethnic backgrounds (Table 7.1), including resections of normal tissue from patients suffering from liver cancer [13], morbid obesity [20], or otherwise healthy, deceased liver donors who may have undergone multiple pharmacotherapeutic treatments immediately prior to death [15,21]. Second, different genotyping and gene expression platforms and analysis methods were used to identify eQTL (Table 7.1). Furthermore, the studies controlled for different covariates and had variable power due to the number of samples, genotypes, and transcripts interrogated (Table 7.1).

Despite these issues, there is evidence that a substantial proportion of the *cis*-eQTL findings are reproducible. Innocenti et al. [21] attempted to replicate findings from their eQTL study in two additional datasets, one of which was a subset of samples from the Schadt et al. study. Sixty-seven percent of the significant *cis*-eQTLs (Bayes factor $(BF) > 5$) replicated in at least one of the additional datasets ($p < 0.05$ and concordant effect direction), but this replication rate dropped to 6% for *trans*-eQTLs ($BF > 5$) [21]. Greenawalt et al. [20] also examined the reproducibility of their liver *cis*-eQTL ($FDR < 10\%$) in the Schadt et al. study, which used the same microarray gene expression platform and found a 66% overlap. Using Bonferroni adjusted p-values, a much more stringent analysis—a comparison between the Schadt et al. and Schroder et al. studies—showed that only 16% of the *cis*-eQTL replicated and there was no replication of *trans*-eQTL [13]. Neither of the latter two studies state whether a concordant effect direction was a criterion for replication, but it appears that liver *cis*-eQTL are comparatively reproducible while *trans*-eQTL are not. This observation may be partly explained by statistical power. Compared with *cis*-eQTL, *trans*-eQTL have a greater multiple test correction burden and thus are less likely to achieve statistical significance in two independent studies.

A salient finding in the Innocenti et al. study [21] was that the *cis*-eSNP with the highest BF in each gene was frequently located immediately upstream or downstream of the transcription start site (TSS) of the eQTL target gene. Additionally, 75% of *cis*-eQTL within 100 kb of the TSS were found to replicate, compared to 61% of those located more than 100 kb from the TSS [21]. Furthermore, Schadt et al. [15] noted that 70% of the *cis*-eQTL were found within 100 kb of the TSS. Therefore, it appears that reproducible liver *cis*-eQTL SNPs are generally found close to gene TSSs. This suggests that eSNPs may directly impact gene transcription by altering gene regulatory regions.

7.2.3 Liver eQTL Tissue Specificity

To examine liver eQTL tissue specificity, Schadt et al. compared their findings with eQTL analyses of human blood and adipose tissue [23]. Approximately 30% of the *cis*-eQTL identified in these tissues were also detected among the liver eQTL (FDR < 10%) from their study [15]. Kabakchiev and Silverberg [24] compared eQTL identified in human ileum, another tissue known to highly express ADME genes, with the liver eQTL identified in the Schadt et al. study [15]; they found that approximately 37% (FDR < 5%) and 43% (FDR < 10%) of significant hepatic *cis*-eQTL were replicated in the intestine. Greenawalt et al. [20] examined replication of their liver *cis*-eQTL in adipose tissue from the same cohort of individuals used for liver eQTL analysis, finding that 46–48% of the liver *cis*-eQTL replicated at $p < 0.05$. Moreover, they found that liver eQTL, but not eQTL identified from adipose tissue, were enriched for the KEGG pathway *drug metabolism–cytochrome P450 genes*. Furthermore, of the 11 ADME genes that had reproducible eQTL between the Schroder et al. and Schadt et al. studies [13,15], *CYP3A5*, *DHRS2*, and *SLC22A3* have been predicted to be preferentially expressed in the liver based on analysis of tissue-specific expression profiles [25]. These findings suggest that at least some fraction of the genetic control of gene expression is specific to the liver.

On the basis of comparisons of eSNPs and gene expression in the liver, subcutaneous and visceral adipose tissue, blood, and skeletal muscle, Fu et al. [26] identified four categories of tissue-dependent *cis*-regulation of gene expression: specific regulation, alternative regulation, different effect size, and opposite allelic direction. It has been demonstrated that rs12740374, which maps to the *SORT1* gene, alters the binding site for C/EBP transcription factors (TFs), consequently modulating hepatic expression of *SORT1* and affecting low-density lipoprotein cholesterol and myocardial infarction risk [27,28]. This eQTL replicated in liver samples but not in blood, subcutaneous and visceral adipose, or skeletal muscle, despite abundant expression of *SORT1* in all tissues, which indicates liver-specific regulation of *SORT1* via this eSNP [26]. In a hepatic example of alternative regulation, rs714885 was found to associate with the regulation of *TMEM176A* (hepatocellular carcinoma-associated antigen 112) expression in the liver; however, an independent, unlinked (LD $r^2 = 0.002$) variant, rs6464104, was found to regulate expression levels of this gene in blood [26]. Even in instances where eQTL and gene associations have not been found to be tissue-specific, variance in both the magnitude and the directionality of effect across tissues has been noted. Specifically, the same eSNP associated with expression levels of D-dopachrome tautomerase (*DDT*) in both blood and liver; however, its effect on expression levels exhibited completely opposite directionality in these tissues. Such findings underscore the complexity of unraveling causative genetic variants as well as the mechanisms by which they mediate tissue-dependent expression, and they reiterate the importance of investigating phenotypicallly relevant tissues.

7.2.4 Correlation of Liver eQTLs and CYP Enzyme Activity

Gene transcript levels are a useful intermediate phenotype to study, but measurements of protein levels or enzyme activity may provide a more informative phenotype for pharmacogenomics. One of the advantages of the liver eQTL model is that subcellular fractions can be generated from the liver, allowing proteomic and biochemical studies to be integrated with genomic and transcriptomic data from the same tissue. For example, Yang et al. [29] examined the enzymatic activities of eight CYP proteins using liver microsomal samples derived from the same cohort used by Schadt et al., [15] and found that the activities of the CYPs were positively correlated with mRNA transcript levels after adjusting transcript and activity levels for age, sex, and study site. Transcript levels explained 10–30% of the variance in enzyme activity, indicating that other biological and maybe environmental or unexplained factors, such as post-translational mechanisms, may also impact activity levels. Fifty-four enzyme activity QTL (aQTL) were identified for the 8 CYP proteins: 24 SNPs were *trans*-aQTL and 30 were *cis*-aQTL (FDR < 10%) [29]. The *cis*-aQTL SNPs were all associated with CYP2D6 enzyme activity, and 29 of these were found to also be *cis*-eSNPs of *CYP2D6* (FDR < 10%). The directionality of the eQTL associations were not reported, however, so the concordance of these associations with enzyme activity cannot be determined.

7.2.5 Web-Based Liver eQTL Data Resources

The Gene Expression Omnibus (GEO) repository contains data from all of the human liver eQTL studies (Table 7.1), although genotypes from the studies carried out by Schadt et al. and Greenawalt et al. are not available. The results of the Schadt et al. study are available through the

TABLE 7.3 Repositories of eQTL Data

Repository	Website	Description
GEO	http://www.ncbi.nlm.nih.gov/geo/	Repository of gene expression data from >2500 studies
SCAN	http://www.scandb.org	Repository of SNP and CNV eQTLs identified in LCLs
eqtl.uchicago.edu	http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/	Genome browser with eQTL annotations
GTEx	http://www.ncbi.nlm.nih.gov/gtex/GTEX2/gtex.cgi	Repository of eQTL data from different tissues

Genotype-Tissue Expression (GTEx) eQTL browser (http://www.ncbi.nlm.nih.gov/gtex/GTEX2/gtex.cgi) and can be queried by gene or SNP and filtered by p-value or LD r^2 values. The GTEx browser provides further useful information by annotating corresponding eSNPs with significant findings from GWAS. While the GTEx eQTL browser currently contains information regarding previous studies, the ongoing Genotype-Tissue Expression (GTEx) project [30] will provide a large-scale database and tissue bank for eQTL analysis across a wide range of human tissue types. The University of Chicago eQTL browser (http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/) provides a graphical display of eQTL identified from numerous studies, including those of Schadt et al. and Innocenti et al., within a user-defined genomic interval, and also provides a means for users to upload their own eQTL annotation for viewing in conjunction with the available published eQTL data. Similarly to the University of Chicago eQTL browser, the seeQTL searchable browser and database [31] employs a graphical user interface to explore eQTL data within genomic intervals; although currently only containing data from brain, monocyte, and LCL eQTL studies, seeQTL plans to incorporate liver eQTL data in the near future. Other GWAS and gene expression repositories as well as additional web-based genomics resources have been extensively reviewed by Glubb et al. [32] Web-based liver eQTL repositories are listed in Table 7.3.

7.3 FUTURE DIRECTIONS

7.3.1 Next-Generation Sequencing

Earlier liver eQTL studies interrogated genetic variation and gene expression at the genome level. Early GWAS studies, in contrast, were based on the assumption that variation in common phenotypic traits is primarily due to common variants. Therefore, SNP coverage was limited to genotyping high-density SNP chips to variants with minor allele frequencies of greater than 5%, and sample sizes for most GWAS studies were underpowered, making them unable to detect the effects of less common and rare variants, despite the fact that both common and rare variants have been identified in many genes affecting drug metabolism and transport [33]. Because of the limitations of high-density SNP chips, the variants identified by GWAS to date are able to explain only a small fraction of the clinical phenotypes examined. Since the minor allele frequencies of uncommon and rare variants are below the threshold of inclusion on commercially available genotyping SNP chips, delineating the effects of these variants on gene expression is best accomplished through sequencing.

Still relatively expensive, NGS nevertheless has advanced to the point where it is becoming feasible to perform whole-genome sequencing of DNA and mRNA to comprehensively characterize genetic variation and the entire transcriptome in a relatively large number of samples. However, sequencing costs for the number of subjects required for a rare-variant GWAS of a complex trait are still prohibitive. Genotype imputation can be performed using data from projects such as HapMap and 1000 Genomes Project to greatly increase the number of genotypes interrogated through high-density SNP chips. In this way, most of the variation accessible through sequencing is captured. Indeed, Innocenti et al. [21] imputed genotypes from HapMap to generate more than 2 million genotypes. Imputation can also be used in conjunction with low-coverage sequencing as a more economical means of genotyping less common and rare variants. However, imputation-based methods, are limited by the depth of available reference panels; variants that are too rare to be represented in a reference panel cannot be imputed.

In addition to the genotyping challenges presented by high-density SNP chips, transcriptome information obtained via microarray technologies has been limited by reliance on existing knowledge about the genome sequence, decreased sensitivity as a result of cross-hybridization of probes and targets, and reduced dynamic range of detection due to background hybridization and signal saturation, making it difficult to detect and/or quantify both high- and low-abundance species. Advances in NGS have led to the use of RNA-seq (next-generation sequencing of mRNA) to define transcriptomes. Pickrell et al. [34] found a high correlation between transcript levels quantitated by RNA-seq and microarrays. Using data from the same 53 individuals, these researchers demonstrated that 70% of SNPs previously identified as eQTL (FDR < 10%) by expression microarrays were replicated by RNA-seq ($p < 0.05$), while 93% of microarray-based associations showed concordant directionality in the RNA-seq data. Montgomery et al. [35] found that approximately 10 million sequencing reads can provide the same depth of coverage as microarrays, but because RNA-seq overcomes many of the limitations presented by microarray technology, improved resolution allows detection of more eQTL. RNA-seq has other advantages over oligonucleotide expression arrays—for example, no saturation of intensity signal for high-abundance transcripts and the ability to quantify allele-specific expression [36]. Additionally, RNA-seq can identify splicing QTL (sQTL), which may not affect overall gene expression but alter transcript isoform levels [34]; and noncoding RNAs (e.g., miRNAs, lncRNAs, etc.), which are known to play a very important role in the regulation of gene expression.

Ongoing efforts by both the Pharmacogenomics Research Network (PGRN) and GTEx are utilizing RNA-seq to create comprehensive public databases with quantitative information on gene expression and regulation in tissues of pharmacological importance. These resources will inform GWAS, provide insights into the mechanisms of gene regulation, and enable future studies of eQTL, AS, and the tissue specificity of gene regulatory mechanisms.

7.3.2 Identification of sQTL

Several thousand liver eQTL that pass multiple comparison criteria have been identified in the studies published to date; however, regulatory mechanisms remain to be elucidated for a vast majority of these associations. Using eQTL data from LCLs, Lalonde et al. [3] estimated that approximately 75% of eQTL regulate gene expression at the transcription level. The remaining 25% are thought to regulate expression levels co- or post-transcriptionally by altering mRNA splicing and/or stability. Veyrieras et al. [37] reported an enrichment of eSNPs in promoter and 3′ regions of genes, and Lalonde et al. [3] found an approximately twofold enrichment of AS events in eQTL targets. For a number of genes, Kwan et al. [38] identified splicing regulatory SNPs (sSNPs) that regulate both expression levels of LCL mRNA and the type of isoforms produced. AS within 5′ and 3′-UTRs may result in the addition or removal of regulatory sequences, thus altering mRNA translation efficiency and/or stability. The selection of an alternative splice site within the 3′-UTR may result in the differential use of stop codons or the use of alternative polyadenylation sites, both of which alter the length and composition of the 3′-UTR.

MicroRNAs (miRNAs) function as post-transcriptional regulators of gene expression by binding to the 3′-UTR of target mRNAs, which generally results in translational repression or mRNA degradation. Therefore, AS that results in altered 3′-UTR length and/or SNPs located in 3′-UTR sequences may eliminate or create miRNA targets, significantly affecting gene expression. Wei et al. [39] identified liver eQTL, or proxies in strong LD ($r^2 = 0.8$), from 12 xenobiotic metabolizing enzymes and transporters that are predicted to create, eliminate, or alter miRNA target sites, suggesting one mechanism by which eQTL might regulate gene expression.

7.3.3 Identification of miRNA eQTL

Modification of miRNA 3′-UTR target sites is only one mechanism through which the regulatory role of miRNAs is coordinated, with genetic variation in miRNA expression levels also likely to have an impact on transcript levels. Indeed, Rieger et al. [40] measured hepatic expression levels of 56 miRNAs predicted to target the 3′-UTR of ADME genes; they found numerous associations with both ADME gene expression profiles and CYP gene expression phenotypes, including mRNA and protein expression, and enzyme activity. Gamazon et al. [41] also found a significant genetic component to miRNA expression variation in the liver, identifying a set of genomic loci associated with miRNA expression levels and enriched for SNPs associated with mRNA expression. More specifically, 28 of 41 VIPs investigated were significantly (FDR < 0.05), and negatively, associated with miRNA expression in the liver. In fact, VIPs were found to be significantly enriched for large miRNA regulatory effects. Given the role of miRNAs in regulating target mRNA expression, miR-eQTL represent genetic variants with significant potential for functional pharmacokinetic and/or pharmacodynamic impact on drug response.

7.3.4 Identification of Functional eSNPs

eQTL analyses are informative, but they cannot definitively identify the causative SNPs that affect gene expression. Because of LD, many eSNPs can represent the same causative genetic variant. Various approaches can be used to identify functional genetic variants that generate the eQTL signal. Reporter gene assays are useful in determining whether eSNPs in promoter, 5′UTR, 3′UTR, and intronic regions have functional effects on gene expression [21], but they cannot be easily applied in a high-throughput fashion and the effects of SNPs in coding regions are generally not amenable to testing.

Bioinformatic analyses may help to narrow the number of putatively causative SNPs by identifying those that are located in functional genomic elements such as are defined by the Encyclopedia of DNA Elements (ENCODE) Project [42]. Among these genomic elements, DNase I hypersensitivity sites (DHS) and a variety of histone modifications have been shown to associate with regulatory regions and mark sites of active transcription and/or repression [43,44]. Indeed, studies of LCL eQTL found that 40–50% of eSNPs were located either in open chromatin regions [44], as evidenced by DNaseI hypersensitivity, or within areas of histone modification.

SNPs in DHS are four times more likely to associate with transcript levels as compared with SNPs not in active chromatin regions [44], while regions of histone modifications known to associate with active promoters and enhancers, but not regions characterized by repressive histone marks, were found to be significantly enriched in eSNPs [44]. Additional analysis of these open chromatin and histone modification regions demonstrated a significant enrichment of TF binding sites in genomic regions containing eSNPs [44], further suggesting that eSNPs have an impact on promoter and/or enhancer function. Moreover, Degner et al., [45] also using LCLs, observed a general correlation between higher DNase I sensitivity and increased TF binding site occupancy, with the genetic variant associating with increased chromatin accessibility also associating with increased gene expression levels for 70% of SNPs that modify both DHS and transcript levels. These findings suggest that epigenetic and other regulatory information can be combined with eQTL data to create a model that ranks eSNPs on their potential to affect gene expression [44]. This approach has not yet been applied to a study of liver eQTL, but could prove extremely useful in identifying pharmacogenetic variants and their target genes.

The mechanism by which distal regulatory elements, such as enhancers and TF binding sites, alter gene expression levels remains poorly understood. However, the relatively recent advent of chromosome conformation capture (3C) technology has led to an appreciation of the important role played by spatial organization of the genome in genetic regulation. The frequency of interaction between any two genomic loci can be determined using 3C and 3C-derived methodologies (e.g., circular chromosome conformation capture, chromosome conformation capture on ChIP, chromosome conformation capture carbon copy, combined chromosome conformation capture ChIP cloning, global 3C interactions, and chromatin interaction analysis by paired-end tag sequencing). These allow detailed quantitative analysis of chromosome structure [46]. Briefly, a fixative agent such as formaldehyde is used to cross-link physically touching segments of interacting chromatin throughout the genome. After proteins that are not cross-linked are solubilized, the DNA in the fixed chromatin is digested using restriction enzymes and religated under conditions that favor intramolecular ligations over ligations between random fragments. The relative frequencies with which different sites become cross-linked are then determined and are proportional to the frequency with which these genomic loci interact.

Analysis of genome-wide interaction frequencies provides information about physical properties and chromosome conformation [46]. A number of 3C studies have demonstrated the formation of chromatin loops that bring putative genomic regulatory elements into contact with target gene promoters [47]. A study by Vakoc et al. [48] directly implicated these looped interactions in gene regulation, demonstrating that disruption of the chromosomal looped conformation results in gene expression abatement. 3C approaches, in combination with other *in vitro* assays, can also inform SNP function. For example, Pomerantz et al. [49] demonstrated that a SNP associated with cancer risk is located in an enhancer sequence that interacts with the promoter of the MYC proto-oncogene and that the SNP increases enhancer activity in a luciferase assay.

Thus, 3C approaches have the ability to identify discrete genetic regions, such as enhancers or silencers, which modulate gene expression; these data can be matched with eQTL studies to determine if eSNPs are located within such regulatory regions. Furthermore, the potential allele-specific effects of eSNPs on interactions between regulatory regions and gene promoters can be examined by sequencing of the interacting region to identify the allelic variant(s) present.

7.3.5 Epistatic eQTL Interactions

eQTL analyses generally examine the effects of only a single genetic locus on gene expression. However, it is possible that interactions of *cis*- and *trans*-eQTL contribute to gene expression regulation. Indeed, in a study of HapMap LCLs, 15% of transcripts were found to be regulated by a *cis-trans* eQTL interaction while in a single-locus analysis the vast majority of the eSNPs were found not to be significantly associated with gene expression [50]. Therefore, important information may be lost by not considering these eQTL interactions, and this approach could be used to detect epistatic pharmacogenetic variants that may not otherwise be identified.

7.3.6 Copy Number Variants

While much emphasis has been placed on the interrogation of SNPs as human liver eQTL, other sources of germline genetic variation in the human genome may also function as liver eQTL. Insertion and deletions (INDELs) in the human genome, typically between 1 bp and 1 kb in length, account for a smaller fraction of genetic variation, and ongoing studies utilizing NGS approaches are building on the current catalog of almost 2,000,000 INDELs [51–55]. CNVs are classified as deletions, duplications, and insertions in the germline genome that are >1 kb in length and, although they are present at a much lower frequency than SNPs and INDELs, they may have a greater impact on phenotypic variation due to the larger fraction of the genome they encompass [53,56–58]. In the context of gene expression, the potential exists for CNVs to alter gene dosage or ablate regulatory regions in the genome; this function has been shown to explain some eQTL signals identified in LCLs [19,59].

Although the role of CNVs as eQTL has not been directly assessed in human liver tissue, pharmacogenetic studies suggest their relevance. The enzyme CYP2D6, predominantly expressed in the human liver and involved in metabolizing up to 25% of all drugs, can be found with more than two gene copies in ~5.5% of Europeans, and leads to a phenotype of increased drug metabolism [60]. One study examined the influence of CNVs in the GST family of detoxifying enzymes in a group of diffuse large B-cell lymphoma patients treated with standard chemotherapeutics. It identified copy number deletions associated with chemotherapy-related toxicities [61]. Recent studies employing genome-wide CNV analyses in LCLs identified numerous CNVs that associate with cytotoxicity for many common anti-cancer therapeutics [62,63]. These studies highlight the potential for a genome-wide assessment of CNVs as eQTL in human liver tissue as a means to enrich our understanding of pharmacogenomics. Through computational approaches to identify CNVs from SNP genotyping microarrays, the opportunity exists to utilize data from existing liver eQTL studies to gain insight into the function of CNVs as human liver eQTL [13,15,21,64].

7.3.7 Liver Protein QTL

Although gene expression levels can often serve as suitable intermediates for drug phenotypes, mRNA levels do not necessarily reflect protein expression or the functional activity of the corresponding enzyme or transporter. However, high-throughput proteome-wide quantification of protein expression has proven much more problematic and less straightforward than transcriptome quantification. Liquid chromatography–tandem mass spectrometry (LC-MS/MS) provides a way forward, but currently the best approach for sensitive and accurate protein quantitation is via targeted proteomic studies [65]. For example, targeted LC-MS/MS analysis of several UGT1A isoforms has been successfully performed in human liver microsomes [66].

Ohtsuki et al. [67] utilized a similar methodology to measure mRNA and protein expression for 12 CYP and 5 UGT enzymes as well as 12 transporters in microsomes or plasma membranes from 17 human liver biopsies. Only the mRNA expression levels of *CYP2B6*, *CYP2C8*, *CYP3A4*, *UGT1A6*, and *SLCO1B1* were highly correlated with protein expression ($r^2 > 0.6$). *CYP3A5*, *CYP3A7*, *CYP2C19*, and *CYP2D6* mRNA levels were moderately correlated ($r^2 = 0.4 - 0.6$), but the remaining enzymes and transporters investigated showed almost no correlation between gene and protein expression levels. Furthermore, only CYP2B6 enzyme activities were more highly correlated with mRNA expression than with protein expression.

Ohtsuki et al.'s study highlights the limited knowledge we have regarding differences in human hepatic protein abundance and reveals the need for protein QTL analyses to characterize the genetic regulation of protein expression. Such studies have been reported in LCL [68,69] and plasma [70,71] and demonstrate the feasibility of this approach. The integration of DNA variation, transcriptome expression, and proteome quantification will enable a more in-depth understanding of gene expression regulation and interindividual variation in pharmacokinetic and pharmacodynamic phenotypes.

7.4 CONCLUSION

There is a strong rationale for using the findings from human liver eQTL studies as a knowledge base in clinical pharmacogenomic studies. Liver eQTL analyses provide insight into the genetic regulation of ADME genes that may not be obtainable from other tissues, and they identify eQTL that are reproducible and unique to the liver. Moreover, liver eQTL data can help explain previously identified pharmacogenomic associations and provide candidates for prospective clinical study.

Pharmacogenetic studies to date have generally been able to explain only a small percentage of drug responses. In contrast, chromatin conformation signatures have the potential to significantly enhance our understanding of gene expression regulation and and our ability to identify the functional genetic variants underlying eQTL signals. Combinations of SNPs, CNVs, and other structural variants may work in concert to produce unique genomic interactions via alternative chromatin conformations that define the gene expression and pharmacogenetic variation seen in individuals [72]. Examination of liver eQTL in this context may lead to the identification of novel biomarkers of disease susceptibility and drug response. In the future, the integration of NGS, bioinformatic, epigenetic, epistatic, CNV, and protein analyses will generate a truly comprehensive map of the genetic regulation of gene expression in the liver. This is certain to profoundly impact the field of pharmacogenomics.

ACKNOWLEDGMENTS

This chapter is an update of a previously published mini-review [73]. For her contributions to the mini-review, the authors would like to acknowledge and thank Neepa Dholakia.

GLOSSARY

Alternative splicing a regulatory mechanism by which multiple mRNA transcripts are generated from a single gene

Cytochrome P450 a large superfamily of heme-thiolate proteins involved in the metabolism of a wide variety of exogenous and endogenous compounds

Copy number variants deletions, duplications, and insertions in the germline genome that are >1 kb in length

Expression quantitative trait loci genetic loci that regulate gene expression

Genome-wide association study (GWAS) analysis of DNA markers across the genome to identify variants associated with particular phenotypes

Linkage disequilibrium nonrandom association of alleles at two or more loci

Single-nucleotide polymorphism a DNA sequence variation in which a single nucleotide is altered

ACRONYMS AND ABBREVIATIONS

3C	chromosome conformation capture
ADME	absorption, distribution, metabolism, excretion
aQTL	activity quantitative trait loci
AS	alternative splicing
BF	Bayes factor
CNV	copy number variant
COMT	catechol O-methyltransferase
CYP	cytochrome P450
dbSNP	single-nucleotide polymorphism database
DHS	DNase I hypersensitivity sites
ENCODE	Encyclopedia of DNA Elements
EBV	Epstein-Barr virus
eQTL	expression quantitative trait loci
eSNP	expression single-nucleotide polymorphism
FDR	false discovery rate
GEO	Gene Expression Omnibus
GSST1	glutathione S-transferase theta 1
GST	glutathione S-transferase
GTEx	genotype tissue expression
GWAS	genome-wide association study
INDELs	insertion and deletions
LC-MS/MS	liquid chromatography–tandem mass spectrometry
LCL	lymphoblastoid cell line
LD	linkage disequilibrium
lncRNA	long noncoding RNA
NCBI	National Center for Biotechnology Information
NGS	next-generation sequencing
miR-eQTL	microRNA quantitative trait loci
miRNA	microRNA
PGRN	Pharmacogenomics Research Network
PharmGKB	Pharmacogenomics Knowledgebase
SNP	single-nucleotide polymorphism
sSNP	splicing single-nucleotide polymorphism
sQTL	splicing quantitative trait loci
TSS	transcription start site
UGT	uridine diphosphate glucuronosyltransferase
UTR	untranslated region
VIP	very important pharmacogene

REFERENCES

[1] Goldstein DB, Need AC, Singh R, Sisodiya SM. Potential genetic causes of heterogeneity of treatment effects. Am J Med 2007;120:S21–5.

[2] Glubb DM, Innocenti F. Mechanisms of genetic regulation in gene expression: examples from drug metabolizing enzymes and transporters. Wiley Interdiscip Rev Syst Biol Med 2011;3:299–313.

[3] Lalonde E, Ha KC, Wang Z, Bemmo A, Kleinman CL, Kwan T, et al. RNA sequencing reveals the role of splicing polymorphisms in regulating human gene expression. Genome Res 2011;21:545–54. http://dx.doi.org/10.1101/gr.111211.110.

[4] Franke L, Jansen RC. EQTL analysis in humans. Methods Mol Biol 2009;573:311–28.

[5] Li H, Deng H. Systems genetics, bioinformatics and eQTL mapping. Genetica 2010;138:915–24.

[6] Skelly DA, Ronald J, Akey JM. Inherited variation in gene expression. Annu Rev Genomics Hum Genet 2009;10:313–32.

[7] Huang RS, Duan S, Shukla SJ, Kistner EO, Clark TA, Chen TX, et al. Identification of genetic variants contributing to cisplatin-induced cytotoxicity by use of a genomewide approach. Am J Hum Genet 2007;81:427–37.

[8] Huang RS, Duan S, Bleibel WK, Kistner EO, Zhang W, Clark TA, et al. A genome-wide approach to identify genetic variants that contribute to etoposide-induced cytotoxicity. Proc Natl Acad Sci USA 2007;104:9758–63.

[9] Huang RS, Duan S, Kistner EO, Bleibel WK, Delaney SM, Fackenthal DL, et al. Genetic variants contributing to daunorubicin-induced cytotoxicity. Cancer Res 2008;68:3161–8.

[10] Choy E, Yelensky R, Bonakdar S, Plenge RM, Saxena R, De Jager PL, et al. Genetic analysis of human traits *in vitro*: drug response and gene expression in lymphoblastoid cell lines. PLoS Genet 2008;4:e1000287.

[11] Wheeler HE, Dolan ME. Lymphoblastoid cell lines in pharmacogenomic discovery and clinical translation. Pharmacogenomics 2012;13:55–70.

[12] Wienkers LC, Heath TG. Predicting *in vivo* drug interactions from *in vitro* drug discovery data. Nat Rev Drug Discov 2005;4:825–33.

[13] Schroder A, Klein K, Winter S, Schwab M, Bonin M, Zell A, et al. Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption, distribution, metabolism and excretion of drugs in human liver. Pharmacogenomics J 2013;13:12–20. http://dx.doi.org/10.1038/tpj.2011.44.

[14] Yu X, Lin J, Zack DJ, Qian J. Computational analysis of tissue-specific combinatorial gene regulation: predicting interaction between transcription factors in human tissues. Nucleic Acids Res 2006;34:4925–36.

[15] Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol 2008;6:e107.

[16] Baune BT, Hohoff C, Berger K, Neumann A, Mortensen S, Roehrs T, et al. Association of the COMT val158met variant with antidepressant treatment response in major depression. Neuropsychopharmacology 2008;33:924–32.

[17] Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. Clin Pharmacol Ther 2012;92:414–7. http://dx.doi.org/10.1038/clpt.2012.96.

[18] Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001;29:308–11.

[19] Gamazon ER, Nicolae DL, Cox NJ. A study of CNVs as trait-associated polymorphisms and as expression quantitative trait loci. PLoS Genet 2011;7:e1001292.

[20] Greenawalt DM, Dobrin R, Chudin E, Hatoum IJ, Suver C, Beaulaurier J, et al. A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. Genome Res 2011;21:1008–16.

[21] Innocenti F, Cooper GM, Stanaway IB, Gamazon ER, Smith JD, Mirkov S, et al. Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. PLoS Genet 2011;7:e1002078.

[22] Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010;6:e1000888.

[23] Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, et al. Genetics of gene expression and its effect on disease. Nature 2008;452:423–8.

[24] Kabakchiev B, Silverberg MS. Expression quantitative trait loci analysis identifies associations between genotype and gene expression in human intestine. Gastroenterology 2013;144:1488–96. http://dx.doi.org/10.1053/j.gastro.2013.03.001. [96.e1-3]

[25] Liu X, Yu X, Zack DJ, Zhu H, Qian J. TiGER: a database for tissue-specific gene expression and regulation. BMC Bioinformatics 2008;9:271. http://dx.doi.org/10.1186/471-2105-9-271.

[26] Fu J, Wolfs MG, Deelen P, Westra HJ, Fehrmann RS, Te Meerman GJ, et al. Unraveling the regulatory mechanisms underlying tissue-dependent genetic variation of gene expression. PLoS Genet 2012;8:e1002431. http://dx.doi.org/10.1371/journal.pgen.

[27] Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43:333–8. http://dx.doi.org/10.1038/ng.784.

[28] Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707–13. http://dx.doi.org/10.1038/nature09270.

[29] Yang X, Zhang B, Molony C, Chudin E, Hao K, Zhu J, et al. Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. Genome Res 2010;20:1020–36.

[30] GTEx Consortium . The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580–5. http://dx.doi.org/10.1038/ng.2653.

[31] Xia K, Shabalin AA, Huang S, Madar V, Zhou YH, Wang W, et al. seeQTL: a searchable database for human eQTLs. Bioinformatics 2012;28:451–2. http://dx.doi.org/10.1093/bioinformatics/btr678.

[32] Glubb DM, Paugh SW, van Schaik RH, Innocenti F. A guide to the current web-based resources in pharmacogenomics. Methods Mol Biol 2013;1015:293–310. http://dx.doi.org/10.1007/978-1-62703-435-7_19.

[33] Solus JF, Arietta BJ, Harris JR, Sexton DP, Steward JQ, McMunn C, et al. Genetic variation in eleven phase I drug metabolism genes in an ethnically diverse population. Pharmacogenomics 2004;5:895–931.

[34] Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature 2010;464:768–72.

[35] Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. Nature 2010;464:773–7.

[36] Majewski J, Pastinen T. The study of eQTL variations by RNA-seq: from SNPs to phenotypes. Trends Genet 2011;27:72–9.

[37] Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, et al. High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet 2008;4:e1000214. http://dx.doi.org/10.1371/journal.pgen.

[38] Kwan T, Benovoy D, Dias C, Gurd S, Provencher C, Beaulieu P, et al. Genome-wide analysis of transcript isoform variation in humans. Nat Genet 2008;40:225–31. http://dx.doi.org/10.1038/ng.2007.57.

[39] Wei R, Yang F, Urban TJ, Li L, Chalasani N, Flockhart DA, et al. Impact of the interaction between 3′-UTR SNPs and microRNA on the expression of human xenobiotic metabolism enzyme and transporter genes. Front Genet 2012;3:248. http://dx.doi.org/10.3389/fgene.2012.00248.

[40] Rieger JK, Klein K, Winter S, Zanger UM. Expression variability of absorption, distribution, metabolism, excretion–related microRNAs in human liver: influence of nongenetic factors and association with gene expression. Drug Metab Dispos 2013;41:1752–62.

[41] Gamazon ER, Innocenti F, Wei R, Wang L, Zhang M, Mirkov S, et al. A genome-wide integrative study of microRNAs in human liver. BMC Genomics 2013;14:395. http://dx.doi.org/10.1186/471-2164-14-395.

[42] Myers RM, Stamatoyannopoulos J, Snyder M, Dunham I, Hardison RC, Bernstein BE, et al. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol 2011;9:e1001046.

[43] Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 2009;459:108–12. http://dx.doi.org/10.1038/nature07829.

[44] Gaffney DJ, Veyrieras JB, Degner JF, Pique-Regi R, Pai AA, Crawford GE, et al. Dissecting the regulatory architecture of gene expression QTLs. Genome Biol 2012;13:R7.

[45] Degner JF, Pai AA, Pique-Regi R, Veyrieras JB, Gaffney DJ, Pickrell JK, et al. DNase I sensitivity QTLs are a major determinant of human expression variation. Nature 2012;482:390–4. http://dx.doi.org/10.1038/nature10808.

[46] Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science 2002;295:1306–11.

[47] Sexton T, Bantignies F, Cavalli G. Genomic interactions: chromatin loops and gene meeting points in transcriptional regulation. Semin Cell Dev Biol 2009;20:849–55. http://dx.doi.org/10.1016/j.semcdb.2009.06.004.

[48] Vakoc CR, Letting DL, Gheldof N, Sawado T, Bender MA, Groudine M, et al. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Mol Cell 2005;17:453–62.

[49] Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet 2009;41:882–4. http://dx.doi.org/10.1038/ng.403.

[50] Becker J, Wendland JR, Haenisch B, Nothen MM, Schumacher J. A systematic eQTL study of cis-trans epistasis in 210 HapMap individuals. Eur J Hum Genet 2012;20:97–101.

[51] Weber JL, David D, Heil J, Fan Y, Zhao C, Marth G. Human diallelic insertion/deletion polymorphisms. Am J Hum Genet 2002;71:854–62.

[52] Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, et al. An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Res 2006;16:1182–90.

[53] Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, et al. Mapping and sequencing of structural variation from eight human genomes. Nature 2008;453:56–64. http://dx.doi.org/10.1038/nature06862.

[54] Mills RE, Pittard WS, Mullaney JM, Farooq U, Creasy TH, Mahurkar AA, et al. Natural genetic variation caused by small insertions and deletions in the human genome. Genome Res 2011;21:830–9. http://dx.doi.org/10.1101/gr.115907.110.

[55] Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061–73. http://dx.doi.org/10.1038/nature09534.

[56] Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nat Genet 2004;36:949–51.

[57] Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature 2006;444:444–54.

[58] Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, et al. Mapping copy number variation by population-scale genome sequencing. Nature 2011;470:59–65. http://dx.doi.org/10.1038/nature09708.

[59] Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 2007;315:848–53.

[60] Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. Pharmacogenomics J 2005;5:6–13.

[61] Cho HJ, Eom HS, Kim HJ, Kim IS, Lee GW, Kong SY. Glutathione-S-transferase genotypes influence the risk of chemotherapy-related toxicities and prognosis in Korean patients with diffuse large B-cell lymphoma. Cancer Genet Cytogenet 2010;198:40–6. http://dx.doi.org/10.1016/j.cancergencyto.2009.12.004.

[62] Kalari KR, Hebbring SJ, Chai HS, Li L, Kocher JP, Wang L, et al. Copy number variation and cytidine analogue cytotoxicity: a genome-wide association approach. BMC Genomics 2010;11:357. http://dx.doi.org/10.1186/471-2164-11-357.

[63] Gamazon ER, Huang RS, Dolan ME, Cox NJ. Copy number polymorphisms and anticancer pharmacogenomics. Genome Biol 2011;12:R46. http://dx.doi.org/10.1186/gb-2011-12-5-r46.

[64] Winchester L, Yau C, Ragoussis J. Comparing CNV detection methods for SNP arrays. Brief Funct Genomic Proteomic 2009;8:353–66. http://dx.doi.org/10.1093/bfgp/elp017.

[65] Kline KG, Finney GL, Wu CC. Quantitative strategies to fuel the merger of discovery and hypothesis-driven shotgun proteomics. Brief Funct Genomic Proteomic 2009;8:114–25.

[66] Harbourt DE, Fallon JK, Ito S, Baba T, Ritter JK, Glish GL, et al. Quantification of human uridine-diphosphate glucuronosyl transferase 1A isoforms in liver, intestine, and kidney using nanobore liquid chromatography-tandem mass spectrometry. Anal Chem 2012;84:98–105.

[67] Ohtsuki S, Schaefer O, Kawakami H, Inoue T, Liehner S, Saito A, et al. Simultaneous absolute protein quantification of transporters, cytochromes P450, and UDP-glucuronosyltransferases as a novel approach for the characterization of individual human liver: comparison with mRNA levels and activities. Drug Metab Dispos 2012;40:83–92. http://dx.doi.org/10.1124/dmd.111.042259.

[68] Wu L, Candille SI, Choi Y, Xie D, Jiang L, Li-Pook-Than J, et al. Variation and genetic control of protein abundance in humans. Nature 2013;499:79–82.

[69] Garge N, Pan H, Rowland MD, Cargile BJ, Zhang X, Cooley PC, et al. Identification of quantitative trait loci underlying proteome variation in human lymphoblastoid cells. Mol Cell Proteomics 2010;9:1383–99.

[70] Johansson Å, Enroth S, Palmblad M, Deelder AM, Bergquist J, Gyllensten U. Identification of genetic variants influencing the human plasma proteome. Proc Natl Acad Sci USA 2013;110:4673–8.

[71] Lourdusamy A, Newhouse S, Lunnon K, Proitsi P, Powell J, Hodges A, et al. Initiative ftAsDN. Identification of cis-regulatory variation influencing protein abundance levels in human plasma. Hum Mol Genet 2012;21:3719–26.

[72] Crutchley JL, Wang XQ, Ferraiuolo MA, Dostie J. Chromatin conformation signatures: ideal human disease biomarkers? Biomark Med 2010;4:611–29. http://dx.doi.org/10.2217/bmm.10.68.

[73] Glubb DM, Dholakia N, Innocenti F. Liver expression quantitative trait loci: a foundation for pharmacogenomic research. Front Genet 2012;3:153. http://dx.doi.org/10.3389/fgene.2012.00153.

This page is intentionally left blank

Chapter 8

Next-Generation Sequencing

Pawel Herzyk
Glasgow Polyomics and The Institute of Molecular Cell and Systems Biology, University of Glasgow, UK

Chapter Outline

8.1 Overview 125
8.2 First-Generation Sequencing: Chain Termination Method 125
8.3 The Concept of Next-Generation Sequencing 126
  8.3.1 Second-Generation Sequencing 126
8.4 Application-Independent Data Analysis 129
  8.4.1 Base Calling 129
  8.4.2 Read Quality Control 129
  8.4.3 Read Assembly 130
  8.4.4 Assembly Through Read Mapping 130
  8.4.5 Sequence or Read Count? 131
8.5 Genomic Applications 131
  8.5.1 De Novo Genome Sequencing 131
  8.5.2 Variant Calling Applications 131
8.6 Transcriptomic Applications 133
  8.6.1 Library Preparation 133
  8.6.2 Sequencing 133
  8.6.3 Transcript Reconstruction 134
  8.6.4 Quantification of Transcript Abundance 135
  8.6.5 Differential Expression 135
8.7 Epigenomic Applications 135
  8.7.1 ChIP-seq 136
  8.7.2 DNA Methylation 136
  8.7.3 Example 137
8.8 Metagenomic Applications 137
  8.8.1 Microbiome Profiling 137
  8.8.2 Metagenomic Sequencing 138
  8.8.3 Example 139
8.9 Conclusions 141
References 141

8.1 OVERVIEW

DNA sequencing is an experimental process of determining the order of nucleotides in a given DNA molecule. Precise knowledge about the DNA sequence of entire genomes is paramount in identifying genes, transcripts, and proteins downstream, and consequently in elucidating biochemical processes taking place in a given organism. Until a decade ago first-generation sequencing was dominated by the chain termination method, developed by Fred Sanger [1], which generated detailed knowledge about the nucleotide sequence of many prokaryotic and several eukaryotic genomes, contributing to various genotyping and mutation discovery projects as well as assessing gene expression levels through serial analysis. The most spectacular achievement, however, was sequencing the entire human genome [2,3], which took a decade at cumulative cost of $2.7 billion [4] before it was finally completed in 2003. The subsequent arrival of next-generation sequencing technologies not only reduced the cost of sequencing and increased throughput but also opened a variety of applications not available to the traditional first-generation sequencing. It is through these applications that new-generation sequencing revolutionized research in genomics, epigenomics, transcriptomics, and metagenomics. In this chapter we discuss the current state of next-generation sequencing technology with a focus on second-generation methods and the available applications.

8.2 FIRST-GENERATION SEQUENCING: CHAIN TERMINATION METHOD

The chain termination method in its contemporary version, which uses automated fluorescent DNA sequencing equipment [5], relies on an in vitro DNA replication reaction, where, alongside normal reagents—single-strand DNA template, DNA primer, DNA polymerase, all four standard deoxynucleotides (dATP, dCTP, dGTP, and dTTP)—a small population of irreversibly blocked, fluorescently labeled dideoxynucleotides (ddATP, ddCTP, ddGTP, and ddTTP) is added to the reaction. The incorporation of a dideoxynucleotide into the newly synthesized chain

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00008-6
© 2014 Elsevier Inc. All rights reserved.

TABLE 8.1 Major NGS Applications

Area	Application Type	Read Assembly	Data Type
Genomic	Sequencing new genomes	De novo	Sequence
	Variant calling	Mapping	Sequence
Transcriptomic	Transcriptome discovery	De novo or mapping	Sequence
	Differential gene expression	Mapping	Read count
Epigenomic	ChIP-seq: protein-binding sites or histone modifications	Mapping	Read count
	DNA methylation	Mapping	Sequence
Metagenomic	Microbiome profiling	Mapping	Sequence
	Metagenomic sequencing	De novo	Sequence

terminates DNA replication because it does not have the required hydroxyl group in the 3′ position. As a result, the final mixture not only contains the population of full-length DNA fragments but, most important, it also contains all possible prematurely terminated subfragments, fluorescently labeled according to the identity of the terminating nucleotides. These subfragments are then ordered according to their size using capillary electrophoresis, a high-resolution separation method, and the final nucleotide sequence is determined from the fluorescent peak trace chromatogram.

Modern DNA sequencers based on Sanger method can run simultaneous electrophoresis in 96 or 384 capillaries, produce up to 1000 base-pair-long reads of very low per-base error of 1 in 100,000. Because of relatively long reads and very high accuracy, this method is still popular for single-pass sequencing of DNA regions shorter than 1 Kbp, which is relevant to the closing of gaps when genomes are sequenced with different technologies.

8.3 THE CONCEPT OF NEXT-GENERATION SEQUENCING

By the time the first draft of the human genome was published in 2001 [2,3], it had become clear that new, faster and cheaper, technologies would be necessary for the more in-depth knowledge needed to analyze genetic variants associated with inherited diseases or to design genome-specific therapies underpinning personalized medicine [6]. Subsequently, the National Human Genome Research Institute (NHGRI) introduced a series of "$1000 genome" grants to stimulate breakthroughs in sequencing technology, aiming to sequence the human genome at the cost of $1000, which led to the emergence of several commercialized methods collectively branded as "next-generation sequencing" (NGS).

Looking from today's perspective and judging by the sheer number of scientific publications that exploit NGS, one can say that the development of NGS technologies has indeed been very successful. Not only did the cost of sequencing become cheaper over recent years, but the whole field has become democratized with the arrival of personal sequencers, which are now affordable to large laboratories or medium-size academic facilities. Furthermore, NGS technologies have enabled many diverse applications to address important questions in genomics, transcriptomics, epigenomics, and metagenomics. Some of the most important applications are listed in Table 8.1 and will be discussed in more detail in later sections. First, however, we will discuss some technological aspects of NGS.

It appears that a generic NGS workflow has three major components: library preparation, sequencing, and subsequent data analysis, where the DNA sequencing component is highly technology-specific while both library preparation and data analysis are more generic and depend on specific applications. Looking at the current landscape of NGS technologies, it seems possible to divide them into second- and third-generation methods, where the former, unlike the latter, are unable to sequence a single copy of DNA molecule and require a library of fragmented and amplified DNA [7].

8.3.1 Second-Generation Sequencing

Second-generation sequencing (SGS) can be described as "massively parallel sequencing of high numbers of spatially separated, amplified, short DNA fragments." SGS is different from the first-generation chain termination method in that the latter is not massively parallel nor are the sequencing fragments spatially separated. Although several SGS technologies have been commercialized to a different level of success, here we will focus specifically on just four of them that currently stand out as the most popular. These are the Roche 454 pyrosequencing technology [8], currently applied in GS FLX Titanium and GS Junior systems; the Illumina "sequencing by synthesis" technology [9,10], applied in Genome Analyser as well as in HiSeq and MiSeq systems; the Life Technologies "sequencing by oligo ligation detection" technology [11], applied in SOLiD NGS systems, and Life Technologies' semiconductor sequencing technology [12], known as Ion Torrent, applied in its Ion Personal Genome Machine (PGM) and Ion Proton systems. Given the SGS definition presented previously, a generic SGS workflow consists of distinct steps, as presented in the following subsections.

8.3.1.1 *DNA Library Preparation*

A DNA library is a population of DNA fragments ready for sequencing. Depending on the application, fragments are either obtained by DNA fragmentation methods such as nebulisation, restriction enzyme digest, or sonication (most applications), or generated by PCR amplification (amplicon sequencing). The fragments are ligated to specific 5′ and 3′ adapters, as all SGS technologies require attachment of the fragments to a solid surface. The adapters provide a universal sequence for hybridization to the oligo-coated surfaces and provide further amplification and sequencing (to be discussed). Each sequencing platform provides its own library preparation kits, which are often tailored to specific applications.

8.3.1.2 *Spatial Separation of Fragments*

In order to sequence the fragment ends in parallel, one needs to separate them spatially. The Roche and Life Technologies platforms use microscopic beads that capture individual single-strand DNA fragments [8,12]. After emulsion amplification (to be discussed), the beads, coated with millions of identical DNA molecules, undergo spatial separation by random distribution either into a two-dimensional matrix of wells able to hold one bead only (Roche 454 and LT Ion Torrent platforms) or onto a glass surface to which they are made to bind covalently (LT SOLiD platform) [11]. Illumina separates single-strand DNA fragments by attaching them in a random fashion to a continuous flat oligo-coated surface provided by a glass flow cell [9,13].

8.3.1.3 *Clonal Amplification*

In order to detect sequencing signals, SGS methods require an *in vitro* cloning step to amplify individual surface-attached DNA fragments, so that sequencing of a single fragment will in fact proceed in a concerted manner on a cluster of identical molecules. The bead-attached fragments (Roche 454, SOLiD, and Ion Torrent) are amplified through an emulsion PCR process [14], where each bead, after capturing a single fragment, is inserted into an emulsion droplet containing all necessary amplification reagents. After the droplets are broken and after subsequent DNA denaturation, beads with millions of single-stranded DNA copies are spatially separated as mentioned previously. The fragments attached to the flat surface (Illumina) are subjected to a bridge amplification process [9,13], which results in the conversion of spatially separated DNA fragments into homogeneous clusters containing millions of identical fragments.

8.3.1.4 *Sequencing DNA Fragments in Parallel*

In this process, the order of the nucleotides in each DNA fragment, represented by homogeneous clusters of fragments situated either on beads or on a flat surface, is determined. The process constitutes the core of different NGS technologies. All four methods discussed here use sequencing-by-synthesis. Upon providing sequencing primers and other method-specific reagents, they monitor the process of incorporation of nucleotides into a newly synthesized DNA chain either via polymerization or ligation reactions. The **Roche-454** and **LT Ion Torrent** platforms both work by quantitatively detecting molecules released during the DNA polymerization reaction, where one pyrophosphate and one hydrogen ion are released for a single nucleotide incorporation. More specifically, Roche-454 technology detects occurrences of pyrophosphate through a cascade of enzymatic reactions where, in the presence of adenosine 5′ phosphosulfate, the ATP sulfurylase quantitatively converts pyrophosphate to ATP, which then acts as a cofactor for luciferase-mediated conversion of luciferin to oxyluciferin. The latter reaction generates light in direct proportion to the amount of ATP. All necessary reagents are available in each pico-litre well, which is connected via optical fiber to a photoelectric matrix able to convert light flashes into electric signals [8]. The LT Ion Torrent semiconductor sequencing technology, on the other hand, focuses on detecting changes in hydrogen ion concentration. The release of hydrogen ions during the DNA polimerization reaction decreases the pH of the solution, and such change is then detected by the ion-sensitive field-effect transistor situated directly under the well and converted directly into an electric signal. The layer of micro-wells is directly connected to the layer of ion sensors to make a single semiconductor chip [12].

Both of the above technologies record an event of nucleotide incorporation without detecting the nucleotide identity. The latter can be determined by a recursive sequence of cycles where only one type of nucleotide is supplied to all wells; the identity of wells showing the positive sequencing signal is recorded and finally the excess of added nucleotides is washed out. The next nucleotide type can then be added to repeat the cycle. In contrast, the other two technologies discussed here determine the identity of the incorporated nucleotides directly. **Illumina** runs the sequencing process in cycles—at the beginning of each cycle all four nucleotides, modified to contain specific cleavable fluorescent dyes as well as reversible terminators, are added. For each computationally determined cluster position, the identity of incorporated nucleotide is then recorded via capturing the emitted fluorescence upon laser excitation. The fluorescent labels are then cleaved away, the incorporated nucleotides deblocked and the free nucleotide excess washed away to enable the next cycle [15]. On the **LT SOLiD** platform the DNA-coated beads, which are covalently bound to the flat surface, are combined with universal primers, ligase enzyme, and a pool of octameric probes. Upon hybridization of a probe complementary to the DNA template in a position adjacent to the primer, the probe is ligated to the primer and the identity of nucleotides in positions 4 and 5 are determined from the fluorescent dye attached to the end of the probe. Upon cleavage of the dye-containing stretch of the last three nucleotides, the system is ready for the next sequencing cycle, which will determine the identity of nucleotides in position 9 and 10. The entire round of ligation-mediated sequencing cycles results in determining the identity of pairs of nucleotides every three nucleotides apart, beginning with positions 4 and 5. The whole process just described is performed five times, where each run is offset from the beginning of the preceding run by one base. The alignment of the five reads results in the full sequence of the DNA fragment, with the advantage that each base has been sequenced twice [11,16].

8.3.1.5 Third-Generation Sequencing

Third-generation methods (TGS), which are beginning to appear, focus on monitoring the natural synthesis of a single DNA molecule without the need for DNA fragmentation and amplification [7]. Library preparation will be greatly simplified and sequencing signals will be collected in real time during the DNA replication process. Although third-generation methods are being actively developed, they have not yet established themselves to the extent that second-generation methods have. However, two technologies have emerged as candidates for leading the field. Pacific Biosciences commercialized their PacBio RS system, which uses sequencing by synthesis and focuses on direct identification of the nucleotides incorporated by DNA polymerase during the DNA replication process. This is not dissimilar to some of the discussed SGS methods. However, two major innovations allow this technology to be categorized as third- rather than second-generation. First, unlike other methods, PacBio RS uses modified nucleotides that have fluorescent labels attached to the terminal phosphate group rather than the base, which allows monitoring of the DNA replication in real time as the fluorophore is cleaved naturally by DNA polymerase. Second, the sequencing takes place on SMRT cells, which are nanofabricated consumable structures patterned with 150,000 zero-mode waveguides of extremely small volume, each holding a single active DNA polimerase molecule immobilized in its detection zone. This enables recording of the sequencing signal from a single nucleotide incorporation using a single DNA template without the need for amplification [17].

Another emerging third-generation technology is nanopore, which uses a single protein channel embedded in an artificial lipid bilayer to monitor the sequence-specific voltage changes during DNA transport through the channel [18]. Although several different companies have been involved in pursuing this approach, it seems that Oxford Nanopore is currently the closest to commercialization [19].

TGS methods promise massive improvements, in comparison to SGS methods, that include higher throughput and faster sequencing speed, leading to faster turnaround, longer reads (at the level of tens of Kbp), small amounts of starting material, and lower cost [7].

8.3.1.6 SGS Sample Preparation

Although NGS technology providers have their own library preparation kits and often their own data analysis tools, here we refrain from describing those and concentrate instead on the generic principles that underpin the use of protocols and methods. Most of the genomic DNA library construction protocols follow a sequence of generic steps, including DNA fragmentation, end repair, adapter ligation, size selection, PCR enrichment, and validation. Most popular DNA fragmentation methods include ultrasonication (e.g., Covaris, Bioruptor) or nebulization. The end repair process consists of a sequence of enzymatic reactions that blunt the ends and provide complementary overhang for ligating adapters. The adapters contain specific sequences allowing for attachment to a provided surface, PCR-based enrichment, initiation of the sequencing process, and barcoding of samples to enable pooling of multiple DNA libraries. After adapter ligation, the library is purified, size-selected to narrow the range of DNA insert length, and amplified to selectively enrich DNA fragments, which contain ligated adapters. The library is then validated by checking the required concentration (e.g., Nanodrop, Qubit) and the library size profile (e.g., Bioanalyzer). Finally, multiple barcoded libraries can be diluted and pooled accordingly.

8.3.1.7 Applications

As we have already mentioned, NGS technology has three major components: library preparation, sequencing, and data analysis. Sequencing is the core of each technology and therefore the most stable, while library preparation protocols and particularly data analysis tools have been developed very dynamically over recent years to cope with the increasing number of NGS applications. This process makes it challenging for sequencing facilities to keep up with the speed of development, which includes introduction of novel or more cost-effective library prep protocols and a plethora of newly developed bioinformatics methods implemented in publicly available or commercial software packages.

There are many NGS applications that are now routinely used. The most common are presented in Table 8.1. As one can see, these can be clustered in several ways depending on a variety of factors. The first factor we discuss is the area of interest in the context of existing "omics." By ***genomic*** applications we understand the applications where biological questions refer to genomic sequence, be it the whole genome, the exome, which can be defined as a part of the genome that is transcribed to mature RNA, or any smaller custom-designed regions representing, for example, a panel of genes related to a particular disease. The questions may relate to sequencing entire new genomes de novo or identifying sequence variation in groups of individual genomes, exomes, or gene panels when compared to the existing reference DNA sequences.

A different family of applications relates to **epigenomics** and answers questions related to changes in the epigenome. Here, we are interested in changes to the function of the genome, rather than genome itself, which are manifested in chromatin structure alteration. This can be elucidated through ChIP sequencing (ChIP-seq) experiments, which combine chromatin immunoprecipitation with NGS to identify a genome-wide landscape of given transcription factor–binding sites or chemical modifications to histone proteins. Similarly, a genome-wide profile of DNA methylation can be determined through, for example, bisulfite sequencing. Contrary to genomic and epigenomic applications, the ***transcriptomic*** family of applications focuses on determining the state of transcriptome, the transcribed genome, to answer questions on genome-wide gene expression patterns, and quantification of differential expression between a number of different conditions, as well as discovery of unknown transcribed sequences, including noncoding RNAs. These can be addressed through RNA sequencing (RNA-seq) experiments where cDNA libraries constructed from total RNA are sequenced. Different library construction protocols allow for interrogation of coding RNA, noncoding RNA, or even micro RNA. ***Metagenomic*** applications concentrate on analyzing entire communities of organisms rather than individuals. By targeted sequencing of a reference gene such as 16SrRNA, followed by detailed analysis of sequences covering its hypervariable regions, one can get information on bacterial community phylogenic profiles. On the other hand, sequencing the entire DNA or RNA complement, known as environmental shotgun sequencing, can provide information on community gene content or gene expression profiles to address questions about biochemical processes going on in the bacterial community colonizing a particular ecological niche.

8.4 APPLICATION-INDEPENDENT DATA ANALYSIS

The initial data analysis steps are common to all applications. These include base calling, quality control, and read assembly. They are discussed next.

8.4.1 Base Calling

Since each of the SGS technologies uses a different method of detection of the sequencing signal, which is measured when a nucleotide is incorporated during the sequencing by synthesis, the process of converting such signal into a nucleotide base, known as base calling, is technology specific. The LT SOLiD is the only technological platform that does not perform base calling and provides the raw data in a *color space* rather than *base space*, where the sequence of colors corresponds to the sequence of tiled di-nucleotides [11]; however, for the sake of clarity we will assume that a generic SGS technology converts the sequencing signal into the *base space*. Furthermore, each technology provides a statistical model that allows for estimating the confidence in a given base call by calculating sequencing error that corresponds to the probability that the base call is incorrect. Such confidence is usually expressed in a per-base quality score, Q, equal to $-10*\log_{10}(\text{error})$ [20]. Consequently, Q30 corresponds to sequencing error of 1 in 1000; Q20 corresponds to the sequencing error of 1 in 100. The base-calling process is usually hidden and performed automatically with a platform-specific software and results in delivery of raw data in the form of read sequences, where each base has an associated quality score. These are usually, but not exclusively, provided in FastQ format.

8.4.2 Read Quality Control

The analysis of millions of raw reads is important for assessing the quality of the sequencing experiment. To this end, one can look, for example, at the percentage of base calls above certain per base quality threshold and assess whether that percentage matches the technology provider specification. The series of box-plots representing quality score distributions per sequencing cycle is also informatory. For some technologies, like Illumina, the quality of base calls is expected to deteriorate throughout the sequencing process, which results in lowering the mean quality value and increase of the variance; however, the speed with which such deterioration takes place reveals whether the experiment can be treated as acceptable or not. In the latter case the data can be rescued by trimming every read to remove a fixed number of bases from the 3′end. As each sample has millions of sequenced fragments it is expected that the base fractions do not change between sequencing cycles and reflect the sample genomic GC content. If this is not the case and a distinguished pattern can be seen on the base-fraction per sequencing cycle plot, this could be indicative of a contamination with a persistent sequence, e.g. sequence adapter. Another QC measure is a distribution of mean read-quality, which one should expect to have a narrow peak, strongly biased towards high Q values, without any multi-modality. Monitoring of a read duplication level reveals whether raw sequences are unique. If they are not it may suggest amplification artefacts during library preparation. All the above QC measurements are usually performed by the manufacturer software, but they can also be provided through standalone software such as FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Even if data quality is as good as suggested by the QC measures mentioned above, there will always be some proportion of "bad" reads, which either have low overall quality, contain adapter sequence or are duplicates, believed to be amplification artefacts. To this end the read filtering process can remove or trim the reads that contain the adapter sequence, e.g. using Scythe software (https://github.com/vsbuffalo/scythe), trim or filter out the reads whose overall quality is lower than a customizable threshold, e.g. using Sickle software (https://github.com/najoshi/sickle) and remove the duplicated sequences using Picard (http://picard.sourceforge.net) or SamTools (http://samtools.sourceforge.net) software. This is particularly important in case of sequencing the unknown genomes, where contaminated or low quality reads may lead to spurious read assembly.

8.4.3 Read Assembly

All existing NGS technologies provide the raw data in form of a vast number of short sequencing reads. Although the length of those reads and their number differ considerably depending on the sequencing technology, the individual reads are meaningless and need to be assembled into longer continuous stretches known as contigs in order to provide some meaning to what actually have been sequenced in a given project. Consequently, the initial part of data analysis of each NGS application, as specified in Table 8.1, is to generate a meaningful assembly that would constitute the solid base for downstream analysis. There are two ways of constructing the read assembly depending on whether the reference sequence is available or not [21]. If it is available, then the entire process is conceptually similar to assembling a jigsaw puzzle when we know what the final picture looks like. We can then map the jigsaw bricks into their proper location on the picture, without giving too much thought into whether the neighbouring bricks fit one another, and this mapping process is at the heart of the puzzle solving. However, if we do not know what the final picture looks like, the process is much more complex as we need to find the bricks that fit one another in a sensible way. This challenge is conceptually similar to the read assembly process in the absence of the reference sequence.

8.4.3.1 De Novo Assembly

The read assembly in the absence of a reference sequence, known as sequencing *de novo*, is typical of projects where genomes or transcriptomes of a given species are sequenced for the first time. At the time when the first-generation, chain termination method was a method of choice for genome sequencing, the *de-novo* read assembly process followed the "overlap-layout-consensus" paradigm [22]. First, reads would be subjected to exhaustive, "all against all" pair-wise sequence comparison, to identify pairs of overlapping reads as measured by fulfilment of some read overlap criteria. In the subsequent layout step, a basic assembly graph would be constructed where nodes correspond to reads and edges are overlaps identified in the previous step. The subgraphs consisting of uniquely overlapped reads would then be collapsed to form new nodes corresponding to contigs, where forking edges would mark contigs not uniquely connected indicating potential repeats. The objective of the layout step, is to simplify the basic assembly graph and to find the optimal assembly by constructing a path that visits each note exactly once, which is very difficult if numerous repeats are present in the graph. In the final process, the identified path is converted to a nucleotide sequence representing the consensus after mapping all contributing reads. The arrival of NGS technologies made the above assembly process very difficult to employ because of a much higher number of reads and far shorter read length. The assembly graphs constructed from the NGS reads became far too complex to resolve during the layout step. To tackle this problem Pevzner and coworkers [23] moved the assembly process from read space to k-mer space, where k-mer were short, fixed-length sequences obtained by hashing the reads (k base pairs long, where usually $k > 19$ and shorter than the read length). The assembly graphs were replaced with de Bruijn graphs where k-mers were nodes connected by edge if they overlapped by $k-1$ base pairs exactly, which means that the pair of adjacent, unique k-mers would be linked by a single edge. These graphs were much easier and quicker to construct, and the repeat structures easily identified. Currently, the most advanced *de-novo* genome assemblers such as Velvet [24], Euler-SR [25], ALLPATH [26] or Abyss [27] all use the de-Bruijn graph concept but differ in the way they handle paired-end information and sequencing errors [21].

8.4.4 Assembly Through Read Mapping

As seen from Table 8.1, most of the NGS applications rely on the knowledge of the reference sequence. The read assembly in the presence of such sequence, known also as read alignment or read mapping, is a process aiming to identify the location in the reference sequence where the particular read originates from. Rather than relying on standard sequence searching programs, which were originally designed to search for sequence homology between evolutionary related sequences, such as BLAST [28], various read-mapping software assume that potential mismatches can come either from species polymorphism or from sequencing errors. Consequently, the criteria for successful read alignment calls would be very stringent, usually allowing no more than a couple of mismatches per read. Although the process of mapping short reads to the reference genome seems conceptually simpler than the process of *de-novo* read assembly, to make such a process accurate and efficient requires complex computational algorithms. The first wave of computer programs, including MAQ [29], SOAP [30] and Mosaik (http://bioinformatics.bc.edu/marthlab/Mosaik/), uses hash-based algorithms, which build hash tables index's, either from the set of input reads or from the reference sequence, and then scan the complementary sequence, reference genome or input reads respectively, against said hash table. This accelerates the alignment by enabling fast access to the information on the location of subsequences in the reference. The second group of programs uses the Burrow Wheeler transform (BWT) concept, where the read sequences are transformed into BWT sequences, which enables a far more efficient way of indexing sequences for the purpose of alignment. The implementations of BWT-based methods in popular programs such as Bowtie2 [31], BWA [32] or Soap2 [33] are over one order of magnitude faster than hash table-based methods at the comparable level of accuracy, although they were reported to be less sensitive than the top hash-based aligners such as Stampy [34]. Moreover, they can use gapped alignment, which is particularly important if longer reads (>75 bp) routinely provided these days by various NGS technologies are aligned.

8.4.5 Sequence or Read Count?

After a read assembly is constructed, either *de-novo* or via read mapping to the reference sequence, we can divide the NGS applications into two groups where downstream analysis is either sequence-based or read-count based. In the case of the former, we actively use the sequence information available in the read assembly, for example, to annotate the newly assembled genome or to identify significant sequence variation in the assembly of multiplexed samples. The above applications stand in visible contrast to applications where downstream analysis relies entirely on the number of reads assigned to the genomic features present in the reference sequence. In the case of transcriptomic applications, such features are genes or transcripts and the corresponding read count relates to the level of gene or transcript expression, while in ChIP-seq applications the read count per fix-length adjacent bins is recorded in order to construct genome-wide profiles for identifying peaks corresponding to protein-binding sites [35].

8.5 GENOMIC APPLICATIONS

Genomic applications can be divided into two main groups: sequencing and annotating new genomes and identifying sequence variants in a number of samples.

8.5.1 De Novo Genome Sequencing

De novo sequencing of large genomes is still a formidable task usually performed by research consortia or very experienced research groups with substantial bioinformatic support. However, the arrival of NGS technologies democratized the field so that de novo sequencing of bacterial genomes can now be done routinely by small research groups equipped with a personal sequencer such as Illumina's miSeq. The cost of sequencing a bacterial genome at, say, 100× depth required for de novo assembly is well under £100 (about $162) per sample, and data analysis tools are easily available. The genome assembly described in the De Novo Assembly section usually leads to generating a number of contigs that may be hard to link given a particular sequencing technology because of repeat regions or sequence biases.

The quality of assembly may be measured with (1) number of contigs (the lower the value the better), (2) N50 value (the higher the value the better), defined as weighted median contig size so that half of the assembly is covered by contigs of size N50 or larger [36]. In the ideal scenario the assembly produces one contig, which covers the entire bacterial genome. The successful assembly is usually a combination of sequencing read length, sequencing depth, library insert length, and choice of assembly tool. The successful assembly must be annotated, a computational process that identifies genes within contigs either by homology searches or de novo gene prediction. It can be carried out using standalone software or, preferably, automated annotating systems such as BASys [37], IMG/ER (https://img.jgi.doe.gov/cgi-bin/er/main.cgi), PGAAP (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/), RaTT [38], and RAST [39]. The typical pipeline uses several HMM (hidden Markov model) based gene predictors as well as BLAST-based homology searches. The identified coding regions are then annotated based on functional resources such as COG clusters (http://www.ncbi.nlm.nih.gov/COG/) or Pfam (http://pfam.sanger.ac.uk/), and downstream metabolic pathway analysis is performed using the KEGG database (http://www.genome.jp/kegg/). However, care should be taken when applying fully automated resources, as manual intervention is often required to improve performance [40].

8.5.2 Variant Calling Applications

Whether performing whole-genome sequencing or targeted sequencing, the techniques for variant calling are the same and rely on the information provided in the SAM/BAM file generated during the read-mapping procedure [41]: the list of alignments including mapped read IDs, their sequences, per-base qualities, mapped genomic coordinates, read directionalities, as well as read-mapping qualities, which relate to the probability of read erroneous mapping. The decision about calling a variant, such as a single-nucleotide polymorphism (SNP) or insertion and deletion (indel), has to be made after careful inspection of each position in the alignment and involves analyzing frequencies of mismatches to the reference and the associated qualities. To this end both the accuracy of the alignment and the accuracy of per-base quality scores are critical to successful variant detection. To reduce the uncertainty associated with variant calling, special care should be taken at various stages of the sequencing experiment and data analysis process [42].

To begin with, one should ensure enough sequencing depth during the sequencing process, as high depth increases confidence in a variant call and is critical for identifying rare variants. At the QC step, a stringent quality-based filtering should be applied, while at the read-mapping step multiple alignments should not be allowed, as both steps reduce the chance of spurious alignment. Prior to the proper variant calling procedure, the alignment data should be post-processed further. To this end, the list of the aligned reads should be sorted according to their genomic coordinates; the read or read pair duplicates, believed to be PCR enrichment artifacts, can be removed within Picard (http://picard.sourceforge.net) or SamTools [41] software. The reads mapped to more than one location should be removed, and the alignment around indels should be reassessed and, as it is likely to be inaccurate, corrected. This can be achieved with GATK software [43]. Last but not least, the quality scores should be re-calibrated to ensure that they reflect the true error rate rather than the error rate estimated based on technological uncertainties [44]. Such recalibration can be achieved with GATK software [43], where calculation of the real sequencing error rate is performed based on alignment stretches free from known SNPs and indels. Although the above preprocessing steps are widely used, at least one study analyzing human exome samples sequenced on the Illumina HiSeq platform found that the impact of data preprocessing on the accuracy of variant calling is much lower then anticipated [45].

The carefully preprocessed data can be used for variant calling or genotype calling. Here, variant calling refers to the identification of insertion, deletion (indel), or SNP at a given alignment position. Similarly, genotype calling refers to a description of a nonreference allele at a given locus in a given individual. Early SNP or genotype calling studies analyzed each sample separately and relied entirely on counting the abundance of alleles at a single site and applying a set of thresholds, for example minimum frequency of the minor allele in a heterozygous genotype. This approach encounters serious problems when applied to cases with low and moderate sequencing depth, as low frequency of one of the heterozygous alleles is likely to happen by chance, resulting in a spurious homozygous call. On the other hand, at low sequencing depth the accumulation of sequencing errors may lead to spurious heterozygous calls. Consequently, at low to medium sequencing depth, the SNP and genotype calls are associated with a high level of uncertainty and the methods based on applying fixed cut-offs are far from accurate because they are unable to quantify such uncertainties [46]. A range of probabilistic methods based mostly on Bayesian statistics have been developed that include popular software such as SamTools [41], GATK [43], or FreeBayes (https://github.com/ekg/freebayes). Here, the genotype calls are made based on posterior probabilities calculated from genotype likelihoods and assigned priors, using Bayesian formula. The genotype likelihood, which represents probability of observing sequencing data at a site given an underlying genotype can be calculated using base call uncertainties captured in recalibrated base qualities. The prior genotype probabilities can be enhanced using information available in SNP databases, such as dbSNP. These can be further improved if many individuals are analyzed together because allele frequencies estimated from merged datasets can be incorporated into priors.

Filtering of variant calls is a standard practice, as the variant caller–generated list is likely to contain a high number of false positives. Such filtering usually checks for deviations from Hardy-Weinberg equilibrium [46] and applies a set of thresholds for minimum read depth, minimum per-base quality, minimum distance from an indel, strand bias, and so forth.

8.5.2.1 Example

Biankin and coworkers [47] performed exome sequencing on the prospectively accrued clinical cohort of 142 patients to define genome aberrations in early sporadic pancreatic ductal adenocarcinoma (PDAC). They ran exome sequencing using both SOLiD and Illumina platforms to identify the genome-wide landscape of aberrations and to identify significantly mutated genes. The analysis of the informative subset of 99 patients with cellularity greater than 20% revealed 2627 high-confidence mutations, of which 2016 were nonsilent. The subsequent analysis of genes with nonsilent mutations that occurred in more than one individual cancer revealed that 16 genes were significantly mutated, half of which were already known (KRAS, TP53, CDKN2A, SMAD4, MLL3, TGFBR2, ARID1A, and SF3B1) and another half were novel mutated genes involved in chromatin modification (EPC1 and ARID2), DNA damage repair (ATM) and other mechanisms (ZIM2, MAP2K4, NALCN, SLC16A4, and MAGEA6). All of these genes except MAGEA6 had attributed confirmed or potential relevance when screened against a number of published experimental biological screens and are therefore good candidates for involvement in somatic driving events of carcinogenesis.

8.6 TRANSCRIPTOMIC APPLICATIONS

In **transcriptomic** studies, the aim is to quantitatively identify the RNA content in samples, usually with the intention of making comparisons between different conditions. Until recently the main technology for the transcriptome analysis was gene expression microarrays, which interrogate the mRNA population by hybridization to an arrayed library of probes. Although microarrays have been very successful, particularly in the analysis of differential gene expression, their use was limited to analyzing gene models that were already known and used for designing the interrogating probes. Conversely, no conclusions could be derived for a part of the analyzed RNA population that would not be captured by a given microarray, and so it can be noted that microarrays did not contribute much to broadening our view of the transcriptome landscape. This situation has changed with the arrival of NGS technology, and particularly its whole–transcriptome-sequencing application known as RNA-seq. Here, the entire RNA population can be interrogated with minimal losses. The reads that can be mapped to the reference genome or assembled de novo contribute to further analysis, but even the reads that fail to map can be investigated further in an attempt to assemble them to completely novel genes. Furthermore, not only can the reads be mapped to the existing transcript models; they also can identify novel splice junctions, putative exons, and noncoding RNAs, and they can be assembled to novel transcripts [48,49]. These make the RNA-seq an extremely powerful tool able to discover new transcriptomic features and consequently able to broaden our view of the transcriptome landscape. Despite the discovery power of RNAseq, it is often used in "conservative" mode when there is no interest in new discoveries, but even then RNAseq has strong advantages over microarrays because it can provide much higher dynamic range, sensitivity, and accuracy [50]. That claim, however, was strongly contested when a rich-content, human transcript array was developed [135].

8.6.1 Library Preparation

In order to perform meaningful RNA-seq experiments, the extracted RNA has to be converted to cDNA and the latter has to be processed further in a fashion similar to genomic DNA processing. The caveat of this approach is that ribosomal RNA (rRNA) which constitutes approximately 90% of total RNA, hinders the coverage of mRNA, usually the focus of the study. One way of dealing with this problem is to select for the "interesting" RNA, and this approach is taken when the poly(A) selection method is applied to capture the mRNA complement in eukaryotic organisms, benefitting from the fact that mRNA 3′ends are polyadenylated. Here, the poly(A)-containing molecules are captured using poly(T) oligo-attached magnetic beads; then the captured mRNA molecules are fragmented and the first strand of cDNA synthesized using reverse transcriptase and random primers followed by second-strand synthesis. Poly(A) selection is a very effective method of enriching mRNA, but it misses nonpolyadenylated mRNAs as well as noncoding RNAs. The other approach used, particularly when RNA is not intact or when polyadenylated species are not the only focus of the study, is rRNA depletion, using methods such as RiboMinus [51] or RiboZero [52]. The resulting rRNA-depleted RNA is enriched for both polyadenylated and nonpolyadenylated RNA, preprocessed RNA, tRNA, and numerous noncoding, regulatory species, which are often species of interest [53]. In such cases, rRNA depletion should be the method of choice; however, one should bear in mind that different library preparation methods may have different sequence coverage, which can introduce biases to mRNA expression quantification [53]. The enriched RNA, either via poly(A) selection or ribosomal depletion, needs be fragmented and converted to cDNA using reverse transcriptase and random primers. These are then processed via standard library construction protocols starting with the end repair step. The fragments are usually 100–250 bp long, although for some applications longer fragments are generated to help with resolving complex transcripts or with de novo transcriptome assembly.

8.6.2 Sequencing

RNAseq experiments are usually performed using short 75–150 bp reads. Although single reads of this length align well to the reference genomes and the resulting read counts provide a reliable measure of exonic expression, the paired-end reads are superior as they help to resolve complex transcripts containing shared exons (to be discussed).

8.6.2.1 Reference Sequence

The RNA-seq approach heavily relies on the existence of an annotated reference sequence given that, in order to quantify expression, one needs not only to assign reads to genes/transcripts/exons but also to calculate their corresponding read counts. If an organism under study has a genome sequenced, the genomic sequence will be available and most of the time the genomic annotations containing positions of genes, transcripts, and contributing exons will be known and available in genome annotation files, usually in GFF/GTF format (http://www.sequenceontology.org/gff3.shtml). The quality of such annotation, however, varies greatly between species and depends on the complexity of the genome and also on the resources put into the annotation process. Usually, as in the case for all major model species, the genome releases, which provide improvements to genomic sequence, are rare but the updates to genome annotation are frequent, as work on annotation improvement is constantly carried out. To this end it is important to realize that for data analysis purposes one needs to record not only the genome version but also the genome annotation release. If RNAseq of a nonmodel species is performed, one can potentially use a reference genome of a relevant model species, hoping that a high proportion of reads will align to the reference despite sequence difference. It has been demonstrated that such an approach works; however, both bias and error increase with the evolutionary distance between model and nonmodel species [54]. The alternative approach, particularly attractive if there is no obvious reference genome available, is to assemble the transcriptome de novo from hundreds of millions of reads pooled across all analyzed samples in a given project. Although at first sight it may seem that such assembly is no different from de novo genome assembly, there are important conceptual differences, which reflect the substantial differences in genome and transcriptome architectures [55].

First, the transcriptome assembly process is expected to end up with a high number of small, disconnected de Bruijn graphs, each one corresponding to a given locus and representing its transcript structure, while the genomic reads can be collapsed into a small number of large graphs representing chromosomes. Second, de novo genome assembly methods rely on the assumption that sequencing coverage (number of reads covering a single nucleotide) is approximately uniformly distributed alongside the entire genome. Consequently, sequence stretches of unusually high coverage are likely to be repeats while unusually low coverage is associated with sequencing artifacts [21]. However, the uniform coverage assumption does not hold for transcriptome assembly. Not only can different transcripts have coverage that differs by several orders of magnitude because of different levels of expression; read coverage may be uneven within a transcript because of sequencing biases. Third, alternative splicing and instances of overlapping loci add to the complexity of de novo transcriptome assembly.

All of these differences resulted in the emergence of transcriptome-specific assembly software such as Trinity [55], Oases [56], and transABySS [57], which has been reviewed and evaluated elsewhere [58]. Regardless of a method of assembly, the new transcripts need to have a functional meaning assigned, which is usually achieved by searching for sequence homology to the annotated transcript and peptide sequences present in various annotated sequence resources.

8.6.2.2 *RNA-Seq Read-Mapping Process*

As we have already mentioned, when discussing genomic read mapping, the standard genomic read aligners allow only near perfect matches. Consequently, when applied to RNA-seq data, the standard aligners do not align reads that span exon-exon junctions because many mismatches can be caused by intronic sequences in the reference genome, leading to biased downstream quantification of gene expression. To deal with this problem one can either extract a collection of transcript sequences from the reference genome and use them as reference for read mapping, or expand the reference genome by complementing it with a set of known exon-exon junctions (Mortazavi et al. [48], Wang et al. [50]). Both strategies, however, greatly limit the scope of RNAseq analyses as they rely entirely on incomplete exon libraries and preclude discovery of novel splice junctions and exons. The alternative, more adventurous, approach is to map the reads to the unannotated genome using splice-aware aligners able to map "junction reads" to the reference genome across splice junctions in the absence of genome annotation. This approach was taken by several groups and resulted in the emergence of several programs, such as TopHat [59], SpliceMap [60], MapSplice [61], GSNAP [62], and QPALMA [63]. Tophat maps the reads in a standard way, constructs putative exon models from the clusters of aligned reads, and puts aside the "unmapped" reads. It then enumerates all possible canonical splice junctions between the putative exon models and selects those that can be aligned to the "unmapped" reads [59]. SpliceMap dissects reads into halves and maps them separately; then it narrows down the possible location of a splice junction and identifies the correct junction using the canonical splice junction definition [60]. MapSplice, GSNAP, and Qpalma map reads allowing for long gaps that can correspond to introns; they then identify the correct splice junction position either (a) by optimization based on the score capturing the quality and diversity of the read alignment around the splice junction [61] or (b) by searching using splice site probabilistic models [62] or models trained using machine learning [63]. In their final identification of splice junctions, both Tophat and SpliceMap use the canonical splice junction definition according to which GT and AG are in the donor and acceptor sites, respectively, GSNAP and Qpalma use splice site models generated using databases of confirmed splice junctions, while MapSplice works without any preconceptions of splice junction architecture.

8.6.3 Transcript Reconstruction

The read alignment step is followed by the transcript reconstruction step where all possible transcripts in a given locus are enumerated. This can be reasonably straightforward if one wishes to rely on the current genome annotation provided within existing GFF/GTF files, or it can be very complex if instead of using incomplete annotation one wishes to assemble transcripts *ab initio* from clusters of aligned and overlapped RNA-seq reads, as proposed in software such as Cufflinks [64] and Scripture [65]. For example, in Cufflinks the aligned fragments, preferably represented as paired-end reads, are assembled within each locus separately, resulting in a collection of loci-specific overlap assembly graphs that are then traversed to resolve and enumerate individual transcripts. Scripture is consistent with Cufflinks in general methodological outline, but differs in algorithmic details.

8.6.4 Quantification of Transcript Abundance

The raw quantification of transcript abundance is done by counting fragments assigned to each transcript. This can become complex in cases when exons are shared between transcripts, as there may be fragments whose assignment to transcripts is ambiguous. In such cases the contribution of a fragment to a transcript may be evaluated using that transcript abundance. For example, when calculating abundances, Cufflinks [64] uses a statistical model where the probability of observing each fragment depends on the abundances of the transcripts the fragment can have been derived from. Since different samples can have a different number of fragments in the library, the raw counts need be normalized to allow meaningful between-sample comparisons. The common practice in early studies was to scale the raw counts by the library size expressed as total number of mapped reads [48,66]. To this end, the normalized transcript abundances are sometimes reported in FPKM units, defined as number of "fragments per kilobase of transcript per million" mapped fragments [64], which is a paired-end extension of an earlier single-end read RPKM concept [48]. FPKM can be considered a raw count scaled by library size and length of the transcript.

It was demonstrated, however, that such global normalization is affected by a small fraction of highly expressed transcripts, which leads to biases in estimates of differential expression [67]. Consequently, several more recent methods corrected the library size scaling factors to take into account that a high proportion of transcripts usually have zero counts; a small proportion of high-count transcripts accounts for a high proportion of counts; and a small proportion of transcripts show differential expression. For example, a popular upper-quartile normalization method [67] that scales the raw count by the upper quartile of all nonzero counts was demonstrated to remove the bias without introducing additional noise. The above methods have been reviewed and their performance evaluated elsewhere on both real and simulated data [68].

8.6.5 Differential Expression

Most of the existing methods for RNA-seq analysis of differential expression are parametric and provide pairwise statistical tests based on the assumption that read counts come from the negative binomial (NB) distribution. In an early RNAseq study, Marioni and coworkers [66] showed that variation among technical replicates can be successfully modeled with Poisson distribution. Later studies showed that modeling variation among biological replicates using a Poisson model leads to overdispersion, and therefore modeling biological variation with NB distribution is justified. Several NB-based methods have been proposed and implemented in differential expression analysis software such as CuffDiff [69], DEseq [70], edgeR [71], and BaySeq [72]. The assumption of NB distribution makes it difficult to deal with multifactorial experiments, as the use of analysis of variance (ANOVA), a standard way of analyzing multifactorial microarray experiments, is prohibited because the assumption of data coming from normal distribution is not fulfilled. Indeed, only few methods able to analyze RNA-seq multifactorial experiments exist, implemented in edgeR [71], DEseq [70], and BaySeq [72] software. The first two are able to analyze general multifactorial experiments using NB-generalized linear models while the third uses an empirical Bayes approach to detect patterns of differential expression across multifactorial experiments.

8.6.5.1 Example

In their classic paper on RNAseq, Mortazavi and coworkers [48] quantified a mouse transcriptome by recording the frequency of gene occurrence in each analyzed sample across three different tissues of brain, liver, and skeletal muscle, using the Illumina GAII platform and single-end 25 bp reads. Despite using poly(A) selection, they managed to dramatically improve the quality of gene models by identifying ~17,000 new features, including ~3000 extended or newly identified 3′ UTRs and ~3000 new or extended 5′ exons as well as ~600 new candidate transcripts. The discovery power of RNAseq was further demonstrated by its uncovering of evidence for 145,000 different splices. The authors introduced the concept of RPKM measure and discussed the observed background level. This early study clearly demonstrated the superiority of the RNAseq approach over microarrays and showed the richness of data that can be generated from simple transcriptomic experiments.

8.7 EPIGENOMIC APPLICATIONS

Epigenomic applications are different from the applications previously discussed. Rather then interrogating genetic information contained in a genome or related to gene expression quantification, here we are interested in chemical modifications of genomic DNA and histone proteins, whose complete record defines the epigenome. For the purpose of this chapter, however, we will also consider genome-wide distribution of binding sites for various DNA-binding proteins as an epigenomic application. Consequently, these applications interrogate not what is in the genome but how the genome is packaged and how accessible it is.

The NGS-based analysis of the distribution of a given histone modification or the distribution of a given protein-binding site makes use of chromatin immunoprecipitation (IP) pull-down to enrich the DNA fraction in question. The DNA is then sequenced using a NGS method and the entire procedure is branded as ChIP-seq [73]. The interrogation of DNA methylation state, on other hand, is often analyzed by combining NGS with bisulphite treatment, which is known as BS-seq [74].

8.7.1 ChIP-seq

To construct ChIP-seq libraries one needs to chemically cross-link DNA to its interacting proteins or the associated histones, and then sonicate the chromatin to short (approximately 100–300 bp) fragments. The "interesting" fragments are then immunoprecipitated using antibodies specific to the DNA-binding protein under consideration or to a specific histone modification before the cross-links are reversed to obtain the DNA fragment population associated with either a given DNA-binding protein or a given modified histone, respectively. Library construction is then completed by ligation of the DNA fragments to the sequencing adapters, enrichment by PCR, and size selection [73]. Usually, for every ChIP-seq library, one should construct a control library by fragmenting DNA but not proceeding to the IP step, proceeding to IP with no antibody, or proceeding to IP with nonspecific antibody such as IgG. Although using control libraries increases the cost of the experiment, it dramatically reduces the number of false positives that appear because of artifacts related to nonuniform fragmentation, copy number variation, nonspecific pull-down, read mapping uncertainty, and GC content–related sequencing bias [75]. The sequencing is usually done with 36–100 bp single-end reads, >5 M reads for a *Drosophila* genome, and >20 M reads per large mammalian genome, as recommended by ENCODE and modENCODE consortia [76].

Data analysis of ChIP-seq experiments relies on mapping the reads to the reference genome and subsequent identification of signals associated with protein-binding sites or attachments of modified histones. Usually, the aligned reads are counted over a number of bins covering the entire genome, and it is determined whether a given bin count is different from the background, where the latter is modeled using Poisson distribution or estimated from the control sample. Different, probabilistic, peak-finding algorithms have been implemented in many available and popular programs, such as Erange [48], FindPeaks [77], MACS [78], QuEST [79], Useq (http://sourceforge.net/projects/useq/), Homer (http://biowhat.ucsd.edu/homer/chipseq/index.html), CisGenome [80], Fseq [81], SISSRs [82], Swembl [83], PeakSeq [84], Qeseq [85], CCAT [86], ChipDiff [87], SICER [88], RSEG [89], TPIC [90], and W-ChIPeaks [91]. The performance of 15 of these algorithms was thoroughly compared, using novel protocols based on validation discriminant analysis using two validation datasets designed for algorithmic benchmarking of key histone modifications [85]. Peaks associated with transcription factor–binding sites are usually narrow while those associated with histones are, on the whole, much wider and their distribution and shape depend greatly on the specific histone modification. Consequently, peak-finding algorithms optimized for detection of TF binding sites do not perform well in identifying histone modification signals and vice versa.

8.7.2 DNA Methylation

Although the ChIP-seq approach can be used to analyze DNA methylation using either methyl-binding domains containing proteins [92,93] or methylated cytosine-specific antibodies [93–95], currently the preferred method for interrogation of methylated DNA fractions is bisulfite sequencing, known as BS-seq [74]. Here, the sequencing library is constructed by subjecting DNA to sodium bisulfite treatment, which converts unmethylated cytosines to uracil with no effect on methylated cytosines. The following PCR amplification converts uracil to thymine, which results in unmethylated cytosines being represented by thymines ($C \rightarrow T$) and methylated cytosines being represented by cytosines ($C^m \rightarrow C$). Consequently, the sequencing of bisulfite-treated libraries results in a clear distinction between methylated and unmethylated cytosines at a single base resolution; however, it does not distinguish between unmethylated cytosines and thymines. This makes data analysis difficult and requires algorithms specifically crafted for the analysis of bisulfite-treated samples. To construct a BS-seq library DNA is initially fragmented and the fragments are ligated to either (a) intermediate adapters or (b) methylated sequencing adapters [35]. The fragments are then denaturated and each strand is subjected to bisulphite treatment followed by PCR amplification. If procedure (a) is selected, sequencing adapters need be ligated to intermediate adapters; however, for procedure (b) the methylated adapters become the proper unmethylated ones.

The amount of required sequencing is much higher than is the case for ChIP-seq because sequencing the entire genome at a reasonable depth using paired-end reads is required. Data analysis is divided into two major steps: read mapping and finding the DNA methylation profile. For several reasons, the alignment of BS-seq to the reference genome is far more difficult than the alignment of untreated reads. First, the searching space is increased because both DNA strands are no longer complementary after the bisulfite treatment. Consequently, a given BS-seq read has an equal chance of coming from either a bisulfite-converted forward strand or a bisulfite-converted reversed strand, as well as from their reversed complements. Second, most of the BS-seq reads have reduced complexity because they are

C-poor. This is due to the fact that methylated cytosines occur nearly exclusively in CpG sequences, which account for less then 2% of di-nucleotides. Thus, most of the reads have C converted to T. Third, mapping reads to the reference sequence is asymmetric because T in BS-seq reads can map to both C and T in the reference sequence, while T in the reference can only map to T in the reads.

Because of these difficulties, the standard DNA read-mapping algorithms are not applicable to the BS-seq read-mapping problem. For this reason, several complex BS-seq mapping algorithms have been designed, implemented in software such as BSMAP [96], BRAT [97], MAQ [29], RMAP [98], BS Seeker [99], BatMeth [100], Bismark [101], seqemehl [102], and B-SOLANA [103]. A comprehensive performance assessment of several of those algorithms has been carried out [104]. Several programs for post-alignment analysis, including Kismeth [105], BISMA [106], and QUMA [107] were developed for locus-specific analysis while GBSA [108] software was developed for genome-wide analysis. The post-alignment GBSA analysis removes duplicate reads assumed to be PCR artifacts, and then calculates coverage at each cytosine site of interest, these including CpG for eukaryotic organisms as well as plant-specific sites such as CHG and CHH, where H can be A, C or T. Methylation signals are then scored at each cytosine site and the cumulative score calculated to analyze transcript features such as promoters, transcription start site regions, and first exon and intron, as well as entire gene bodies. Lastly, methylation domains can be identified in an annotation-independent manner using a scoring algorithm based on scanning the entire genome with a sliding window. The domains are then annotated using proximal genomic features available in genome annotation files.

8.7.3 Example

A good example of a powerful epigenomic study is the pioneering work of Barski and coworkers [73], who used Illumina-based SGS to generate a high-resolution map for the genome-wide distribution of 20 histone lysine and arginine methylations in the human genome. They identified the position and shape of each histone modification signal and showed that active promoters are characterized by high levels of H3K27me1 (single methylation at lysine 27 of histone 3), H3K36me3, H3K9me1, H4K20me1, and H2BK5me1 downstream of transcription start sites (TSS). This pattern can potentially be used to identify novel transcriptional units and to confirm TSSs.

By correlating histone modification patterns with gene expression data, Barski and coworkers found out that H3K27me2 and H3K27me3, respectively, correlate significantly with gene silencing. Moreover, they found that H3K9me3 and me2 have significant correlation albeit to much lesser extent. They also confirmed that single, double, and triple methylations of H3K4 correlate well with gene activation, as do H3K27me1, H3K9me1, H4K20me1, and H2BK5me1, while both H3R2me1 and H3R2me2 show a corresponding modest effect. Among other things, their data clearly showed that in the case of H3K79 single methylation is modestly associated with gene activation, triple activation is associated with gene repression, and double methylation shows no preference. The data also clearly demonstrated that ChIPseq is a very powerful method that provides far superior high-resolution data at a lower cost when compared to the traditional ChIP-on-chip approach, which relies on identification of the “chipped” fragments by hybridization to microarray rather than sequencing.

8.8 METAGENOMIC APPLICATIONS

All of the applications discussed so far apply in situations where the sequencing experiment is done using samples from separated known species. This implicit assumption allows us to assemble reads de novo to construct the unknown genome or to map them onto the known single-reference genome. Here, we will discuss a new situation where the sequencing experiment is performed using DNA harvested not from a single species but from the entire microbial community inhabiting a given ecological niche. NGS studies on microbial communities revolutionized and democratized microbial research and increased our understanding of microbial communities.

It has been demonstrated that there are ten times more bacterial cells in the human body than there are cells making up that body [109,110], which leads to the logical conclusion that microbial communities inhabiting the human body must have a pronounced effect on various human conditions. Indeed, altered microbial compositions in the human body have been linked to conditions such as obesity, vaginosis, caries, and inflamatory bowel disease [111–115], and this linkage has led to major research collaborative projects like the Human Microbiome Project and MetaHit whose aim is to analyze the composition and nature of various microbiota, as well as their variability within and between different subjects. Here, we will focus on two NGS applications: microbiome profiling by 16SrRNA gene sequencing and proper metagenomic sequencing of the microbial community metagenome.

8.8.1 Microbiome Profiling

Next-generation sequencing applied to targeted amplicon studies focuses on one or a few phylogenetic marker genes and provides a cost-effective culture-independent method to study the composition and diversity of complex microbial communities. The most widely used gene is 16SrRNA, an excellent phylogenetic marker because it consists of nine different variable regions flanked by conserved stretches where the former are used as markers while the latter facilitate the design of primers for PCR amplification of the variable regions. Several studies have advocated using different variable regions either separately or in tandem [116–118], and various primer pairs of different levels of specificity have been proposed [119,120]. The resulting amplicons are sequenced using NGS technologies of relatively long reads. After the initial filtering step, the sequenced reads can be used either to identify the composition of the microbial community through classification methods or to explore microbial diversity via phylogenetic clustering. For identification, reads can be assigned to different taxonomic units using the Ribosomal Database Project Classifier [121], a naive Bayesian classifier developed for classification of bacterial rRNA sequences into the Bergey bacterial taxonomy [122], without the need for sequence alignment. To explore diversity, the read multiple sequence alignment is usually constructed first, followed by read clustering into operational taxonomic units (OTU) that can be assigned to known taxonomic groups. Phylogenetic trees can then be inferred followed by the analysis of diversity within or between samples.

Pyrosequencing using the Roche 454 platform was initially the most successful because it provided long reads (longer than 400 bp) that could generate high-resolution results when identifying taxa at the species level. However, the pyrosequencing complex error model required strong noise reduction in order to decrease the number of false OTUs [123]. The GAIIx and HiSeq Illumina sequencers entered the race, but deteriorating read quality on the 3′ end was a source of poor accuracy and classification efficiency [117]. Nevertheless, it seems that new developments in sequencing chemistry for the Illumina miSeq and Life Technologies PGM platforms will allow sequencing of 2 × 300 bp reads and single-end 400 bp reads, respectively, which promises to shift microbial profiling toward low-cost personal sequencers.

8.8.2 Metagenomic Sequencing

Metagenomics studies the collective genome of a community of uncultured microorganisms. Rather than use a targeted amplicon approach, metagenomic studies focus on sequencing the entire DNA complement of a given microbial community. To this end these studies concentrate not on composition and diversity but rather on gene content and the functional potential of genes given a particular ecological niche.

During sample preparation, care should be taken so that extracted DNA is representative of all cells in the samples and that potential DNA contamination is kept to a minimum. This is particularly important if the analyzed community is associated with the host of a large genome, as the host's genetic material may be sequenced unnecessarily. Experimental techniques like fractionation and selective lysis are used to minimize host material contributions to the sample [124,125]. Sequencing is carried out using the same approach as for whole genome sequencing of a single species, and subsequently the metagenomic reads need to be assembled to contigs that are then subjected to downstream analysis, leading to functional annotation. As in the case of a single-species sample, assembly may be done by mapping reads to the reference genome or by assembling them de novo. Although mapping metagenomic reads to the reference sequence using software such as Newbler (Roche), Amos (http://sourceforge.net/projects/amos/), or Mira [126] is effective, it only works for the small proportion of reads that come from species whose genomes can be included in the reference sequence. More often than not, however, metagenomic projects use de novo assembly because constructing relevant reference sequences would be greatly incomplete and would seriously limit the scope of the analysis.

Prior to de novo assembly, it is important to clean the data from all known contaminating sequences by running sequence similarity searches—for example, removing host eukaryotic sequences [127]. The main difficulty with assembling metagenomic reads is that the reads are derived from a high number of different microbial species. Contigs generated with standard de Bruijn graph assemblers are usually prematurely terminated—rare *k*-mers generated from reads derived from less abundant species are often removed under the assumption that they represent sequencing errors. This will negatively affect the accuracy of downstream functional analysis. Common regions in different genomes further complicate assembly [128]. New de Bruijn graph de novo assemblers, such as MetaVelvet [129] and Meta-IDBA [130], have been generated specifically for metagenomic applications with the difficulties just mentioned in mind. Both produce significantly longer contigs and are able to reconstruct sequences for low-abundance species.

In a process known as binning, the assembled contigs are sorted into groups likely to represent a single genome or closely related genomes. This is done by clustering contigs based on their GC content and *k*-mers abundance distribution. The resulting genomes, represented by groups of co-binned contigs, are then functionally annotated [128]. If large contigs are produced, standard web-based genome annotation software developed for bacterial genome annotations, such as BASys [37] and RAST [39] can be used. Their main advantage is that they can perform many different analyses, such as homology searches and gene prediction, in a fully automated fashion. Shorter contigs can be searched for the presence of genes, by sequence homology searches (e.g., BLAST) against existing protein sequence databases or, failing that, using tools for *ab initio* gene prediction [127]. The latter approach, usually based on pattern recognition algorithms or supervised learning, is challenging because many open reading frames are partial and therefore may not be easily recognized by *ab initio* predictors. An excellent tool for the analysis is a web-based server known as MG-RAST [131], which accepts collections of assembled contigs or raw reads for providing the functional annotations.

8.8.3 Example

In a recent study which used metagenomic sequencing to investigate microbiota and its functional potential Wang and coworkers analyzed 16 samples collected from dental swabs and plaques representing four periodontal states [132]. The sequencing used the Illumina HiSeq2000 platform providing 2 × 100 bp paired-end reads. After screening out human DNA contamination and quality filtering, the paired-end reads were aligned into 140–194 bp fragments showing between-reads overlap in the range of 6–60 bp. Identification of 16S rRNA sequences was performed using a RDP classifier. The analysis identified 34 novel oral genera present in three or more individual samples. Moreover, a strong correlation between community structure and disease status was observed. Subsequently, the metagenomic read assembly generated contigs that were further functionally annotated using the KEGG database. Wang and colleagues' analysis revealed overrepresentation of functional genes and metabolic pathways in the microbiomes in periodontal disease, related mainly to bacterial chemotaxis, flagellar assembly, and toxin biosynthesis. The researchers suggested that periodontal microbes can adapt to subgingival niches by overrepresentation of genes that may facilitate microbial location, colonization, and invasion of host tissues.

BOX 8.1 Targeted Sequencing

More often than not in NGS applications, researchers are interested in sequencing of certain DNA regions rather then whole genomes. To generate the final DNA library, the genomic DNA must be reduced to include only the regions of interest, which increases the cost-effectiveness of a given project. One would argue that sample preparation for ChIPseq and RNAseq is a good example of genome/transcriptome reduction, as those applications use immunoprecipitation pullouts or poly(A) selection and ribosomal reduction, respectively, to retain the interesting DNA fraction and remove the unwanted sequences.

Here, however, we discuss approaches to targeted sequencing, known as sequence capture, that provide "target enrichment" of interesting genomic subregions by precise design based on knowledge of the whole genome sequence. Targeted sequencing approaches are particularly important where relatively small genomic subregions must be analyzed in a high number of samples. In this case the low cost of sequencing can leverage the high cost of sequence capture.

Several target enrichment methods have been proposed with the aim of maximizing capture accuracy, specificity, and completeness, and minimizing cost and sample preparation time. These can be grouped into three categories: hybridization capture, multiplex PCR amplification, and selective target circularization [133,134]. **Hybridization capture** relies on generating a library of hybridization probes designed to complement the targeted region of interest. These probes are then used as baits against the fragmented library generated from a genome sample in a given NGS technology–specific fashion by ligating specific sequencing and multiplexing adapters to enable multisample pooling. Hybridization can be done on a microarray, with the probes chemically attached to a solid flat surface, or in solution, with the probes attached to magnetic beads. After hybridization is completed, the excess material is washed away and the harvested genomic fragments are recovered in a process specific to a particular enrichment kit. "In solution" methods have gained popularity over recent years because of their lower cost and simplified workflow. Hybridization capture methods are easily scalable and demonstrate a high level of uniformity, which makes them particularly successful for targeting whole exomes. Ready-to-use solutions are provided by several vendors, the most popular being the SureSelect range from Agilent, SeqCap EZ range from NimbleGen, TruSeq Exome Enrichment Kit from Illumina, and Ion TargetSeq Exome Kit from Life Technologies.

Selective target circularization methods are conceptually different although the target fragments are enriched by hybridization to the library of probes. Here, however, the probes, known as molecular inversion probes, padlock probes, or selector probes, are composed of two oligos, complementary to two noncontiguous stretches of a target region, and a DNA linker attached to them in the opposite linear order. Circular molecules are generated by gap-filling ligation chemistry, which ensures that only perfectly hybridized fragments are circularized. These fragments are then amplified and finally the linear fragments are excised from the circular constructs. The latter are ready for sequencing provided that the DNA linker contain NGS method–specific sequencing adapters and sample barcodes. Selective target circularization methods are more specific than hybridization capture methods, but they are less scalable. The HaloPlex range ready-to-use solution from Agilent is an example of a very popular enrichment kit.

In **PCR enrichment** many PCR reactions are carried out to generate amplicons that cover targeted regions. Although conceptually very simple, this approach faces serious challenges because a very high number of PCR reactions usually have to be done in parallel. Several vendors offer ready-to-use kits for long-range PCR amplification, product purification, and normalization, such as SeqTarget from Qiagen. Multiplexed PCR kits allowing many primer pairs to be run in a single multiplexed reaction are also available—for example, Ion AmpliSeq by Life Technologies. High-throughput parallelized systems are available from Fluidigm (http://www.fluidigm.com/targeted resequencing.html) and Raindance (http://raindancetech.com/targeted-dna-sequencing/).

Targeted enrichment approaches have been used successfully in a variety of applications designed for enrichment of different targets. These applications may be divided into three groups, exome enrichment, enrichment of predefined gene panels, and custom enrichment.

BOX 8.2 Exome Enrichment Kits

Several solutions are available for exome enrichment, most of them based on hybridization capture. They include SureSelect from Agilent and SeqCap EZ from NimbleGen. Both are available in two forms, one that enriches human coding regions only and one that enriches both coding regions and UTRs. Other options include TruSeq Exome Enrichment Kit from Illumina and Ion TargetSeq Exome Kit from Life Technologies, both of which are comprehensive solutions for enrichment coding regions, UTRs and various noncoding genes; they are specific to Illumina and Ion Torrent NGS technologies, respectively. CDS-only exome enrichment solutions based on selective circularization or multiplexed PCR enrichment are also available. They include Haloplex Exome Kit from Agilent and Ion Ampliseq Exome Kit from Life Technologies. Both significantly reduce sample preparation time compared to hybridization capture. Exome sequencing has become very popular and contributes significantly to identification of the causative variants for Mandelian diseases.

BOX 8.3 Gene Panel Enrichment Kits

Several vendors offer enrichment kits for disease-specific gene panels. Currently, the most comprehensive of these come from Life Technologies, which provides several ready-to-use panels for Ion Torrent NGS based on their Ion Ampliseq^{TM} target selection technology, which uses multiplexed PCR enrichment. These include the Ion AmpliSeq Cancer Hotspot Panel, which provides hotspot regions including about 2800 COSMIC mutations in 50 oncogenes and tumor suppressor genes; the Ion AmpliSeq Comprehensive Cancer Panel, which provides exons in more than 400 oncogenes and tumor suppressor genes; and Ion AmpliSeq Inherited Disease Panel, which enriches more than 300 genes associated with more than 700 inherited diseases. These panels require between 10 and 40 ng of DNA, which makes them applicable to formalin-fixed, parafin-embedded (FFPE) samples.

Life Technologies also offers Ion Ampliseq Community Panels, which are designed by a consortia of experts working in a given field to provide either comprehensive coverage of mutations in notoriously difficult genes or comprehensive panels of genes related to a particular disease. Currently, two community panels are available: Ion Ampliseq BRCA1 and BRCA2 Panel, covering mutations in the two tumor suppressors implicated in breast and ovarian cancers; and Ion AmpliSeq Colon and Lung Cancer Panel, covering hot spots in 22 genes implicated in colon and lung cancers.

Four other community panels are currently under development. These are the Ion AmpliSeq AML Panel, which will enrich the coding regions of 21 genes implicated in acute myeloid leukemia; the Ion Ampliseq CFTR Panel, which will cover all exons, intron-exon boundaries, and UTRs of the cystic fibrosis transmembrane regulator (CFTR) gene harboring hundreds of mutations implicated in cystic fibrosis; the Ion AmpliSeq Cardio Panel, which will cover all exons and UTRs of 62 genes implicated in cardiomyopathies, channelopathies/arrhythmias, and structural heart defects; and the Ion AmpliSeq TP53 Panel, with will cover all exons and UTRs of the tumor suppressor TP53 gene, which is implicated in many human cancers and is notoriously difficult for probe design because of the presence of repetitive sequences. Of other vendors, Roche Nimblegen provides the hybridization capture–based Neurology Panel Design, which covers mutations in 256 genes associated with 87 neurological disorders; and the Comprehensive Cancer Design panel, which covers mutations in 578 cancer-implicated genes. Illumina provides the TruSeq Amplicon Cancer Panel, which covers mutations in 48 cancer-related genes.

BOX 8.4 Custom Enrichment Kits

Many vendors provide custom enrichment solutions where probes (hybridization capture and selective target circularization) or primer pairs (PCR enrichment) are designed for customized sets of genes and cover CDSs only or both CDSs and UTRs. They can be ordered and delivered as custom gene panels. Such solutions rely on web-based design tools such as SureDesign (Agilent, https://earray.chem.agilent.com/suredesign/), which creates probes for both SureSelect and HaloPlex target enrichments; the NimbleDesign tool (Roche NimbleGen, https://design.nimblegen.com/), which custom-designs SeqCap EZ Choice Libraries; the Ion AmpliSeq Designer tool (Life Technologies, https://www.ampliseq.com/), which designs Ion AmpliSeq Custom Panels; and DesignStudio (Illumina, http://www.illumina.com/applications/designstudio.ilmn), which designs probes for the TruSeq Custom Enrichment kit.

8.9 CONCLUSIONS

The variety of applications presented here clearly demonstrate the power and the breadth of next-generation sequencing technologies. The limitations of second-generation sequencing, however, impact read length and accuracy as well as library preparation workflows. It seems that widely anticipated third-generation methods may overcome at least some of these and generate far longer reads from a single DNA molecule. This would improve the read assembly generated either de novo or via read mapping and consequently affect the quality of downstream analyses as well as facilitate studies on structural genomic variants.

Second-generation sequencing not only revolutionized genomic research but also had a strong impact on development of other "omics" such as proteomics and metabolomics. To this end, it seems that integration of data generated through various "omics" technologies will pose a strong challenge for the future.

REFERENCES

[1] Sanger F, Nicklen S, Coulson R. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 1977;74(12):5463–7.

[2] Lander ES. Initial sequencing and analysis of the human genome. Nature 2001;409(6822):860–921. http://dx.doi.org/10.1038/35057062/.

[3] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science 2001;291(5507):1304–51. <http://www.sciencemag.org/content/291/5507/1304.abstract>.

[4] Pareek CS, Smoczynski R, Tretyn A. Sequencing technologies and genome sequencing. J Appl Genet 2011;52(4):413–35. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3189340&tool=pmcentrez&rendertype=abstract> [cited 07.08.13].

[5] Edwards A, Voss H, Rice P, Civitello A, Stegemann J, Schwager C, et al. Automated DNA sequencing of the human HPRT locus. Genomics 1990;6(4):593–608. <http://www.sciencedirect.com/science/article/pii/088875439090493E>.

[6] Collins FS, Green ED, Guttmacher AE, Guyer MS. A vision for the future of genomics research. Nature 2003;422(6934):835–47.

[7] Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. Hum Mol Genet 2010;19(R2):R227–40. <http://www.ncbi.nlm.nih.gov/pubmed/20858600> [cited 18.09.13].

[8] Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in genome sequencing in microfabricated high-density picolitre reactors. Nature 2005;437(7075):376–80.

[9] Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. Nucleic Acids Res 2006;34(3):e22. <http://nar.oxfordjournals.org/content/34/3/e22.abstract>.

[10] Turcatti G, Romieu A, Fedurco M, Tairi A-P. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. Nucleic Acids Res 2008;36(4):e25. <http://nar.oxfordjournals.org/content/36/4/e25.abstract>.

[11] Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, et al. Accurate multiplex polony sequencing of an evolved bacterial genome. Science 2005;309(5741):1728–32. <http://www.sciencemag.org/content/309/5741/1728.abstract>.

[12] Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. Nature 2011;475(7356):348–52. http://dx.doi.org/10.1038/nature10242.

[13] Adessi C, Matton G, Ayala G, Turcatti G, Mermod J, Mayer P, et al. Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. Nucleic Acids Res 2000;28(20).

[14] Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. Proc Natl Acad Sci 2003;100(15):8817–22. <http://www.pnas.org/content/100/15/8817.abstract>.

[15] Bennett S. Solexa Ltd. Pharmacogenomics. Future Medicine 2004;5(4):433–8. http://dx.doi.org/10.1517/14622416.5.4.433.

[16] Housby JN, Southern EM. Fidelity of DNA ligation: a novel experimental approach based on the polymerisation of libraries of oligonucleotides. Nucleic Acids Res 1998;26(18):4259–66. <http://nar.oxfordjournals.org/content/26/18/4259.abstract>.

[17] Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. Science 2009;323(January):133–8.

[18] Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, et al. The potential and challenges of nanopore sequencing. Nat Biotechnol 2008;26(10):1146–53.

[19] Maitra RD, Kim J, Dunbar WB. Recent advances in nanopore sequencing. Electrophoresis 2012;33(23):3418–28. <http://www.ncbi.nlm.nih.gov/pubmed/23138639> [cited 15.08.13].

[20] Ewing B, Hiller L, Wendl MC, Gree P. Base-calling of automated sequencer traces using Phred I. Accuracy assessment. Genome Res 1998;8:175–85.

[21] Flicek P, Birney E. Sense from sequence reads: methods for alignment and assembly. Nat Methods 2009;6(11):S6–12.

[22] Bonfield JK, Smith KF, Staden R. A new DNA sequence assembly program. Nucleic Acids Res 1995;23:4992–9.

[23] Pevzner PA, Tang H, Waterman MS. An Eulerian path approach to DNA fragment assembly. Proc Natl Acad Sci USA 2001;98(17):9748–53. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=55524&tool=pmcentrez&rendertype=abstract>.

[24] Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008;18(5):821–9. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2336801&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[25] Chaisson MJ, Pevzner PA. Short read fragment assembly of bacterial genomes. Genome Res 2008;18:324–30.

[26] Butler J, Maccallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, et al. ALLPATHS: De novo assembly of whole-genome shotgun microreads. Genome Res 2008;18:810–20.

[27] Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM. ABySS: A parallel assembler for short read sequence data ABySS: A parallel assembler for short read sequence data. Genome Res 2009;19:1117–23.

[28] Korf I. Serial BLAST searching. Bioinformatics 2003;19:1492–6.

[29] Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 2008;18:1851–8.

[30] Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. Bioinformatics 2008;24(5):713–4. <http://www.ncbi.nlm.nih.gov/pubmed/18227114> [cited 07.08.13].

[31] Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10(3):R25. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2690996&tool=pmcentrez&rendertype=abstract> [cited 07.08.13].

[32] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25(14):1754–60. <http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btp324>.

[33] Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, et al. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 2009;25:1966–7.

[34] Lunter G, Goodson M. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res 2011;21(6):936–9. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3106326&tool=pmcentrez&rendertype=abstract> [cited 13.08.13].

[35] Simon SA, Zhai J, Nandety RS, McCormick KP, Zeng J, Mejia D, et al. Short-read sequencing technologies for transcriptional analyses. Annu Rev Plant Biol 2009;60:305–33. <http://www.ncbi.nlm.nih.gov/pubmed/19575585>.

[36] Salzberg SL, Yorke JA. Beware of mis-assembled genomes. Bioinformatics 2005;21(24):4320–1. <http://www.ncbi.nlm.nih.gov/pubmed/16332717> [cited 01.10.13].

[37] Van Domselaar GH, Stothard P, Shrivastava S, Cruz JA, Guo A, Dong X, et al. BASys: a web server for automated bacterial genome annotation. Nucleic Acids Res 2005;33(Web Server issue):W455–9. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1160269&tool=pmcentrez&rendertype=abstract> [cited 25.09.13].

[38] Otto TD, Dillon GP, Degrave WS, Berriman M. RATT: rapid annotation transfer tool. Nucleic Acids Res 2011;39(9):e57. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3089447&tool=pmcentrez&rendertype=abstract> [cited 23.09.13].

[39] Aziz R, Bartels D, Best A, DeJongh M, Dis T, Edwards R, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 2008;9:75.

[40] Kisand V, Lettieri T. Genome sequencing of bacteria: sequencing, de novo assembly and rapid analysis using open source tools. BMC Genomics 2013;14(1):211. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3618134&tool=pmcentrez&rendertype=abstract> [cited 19.09.13].

[41] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25(16):2078–9. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2723002&tool=pmcentrez&rendertype=abstract> [cited 19.09.13].

[42] Altmann A, Weber P, Bader D, Preuss M, Binder EB, Müller-Myhsok B. A beginners guide to SNP calling from high-throughput DNA-sequencing data. Hum Genet 2012;131(10):1541–54. <http://www.ncbi.nlm.nih.gov/pubmed/22886560>

[43] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–303.

[44] Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K. SNP detection for massively parallel whole-genome resequencing. Genome Res 2009;19:1124–32.

[45] Liu Q, Guo Y, Li J, Long J, Zhang B, Shyr Y. Steps to ensure accuracy in genotype and SNP calling from Illumina sequencing data. BMC Genomics 2012;13(Suppl. 8):S8. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3535703&tool=pmcentrez&rendertype=abstract> [cited 18.09.13].

[46] Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet 2011;12(6):443–51. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3593722&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[47] Biankin AV, Wadell N. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 2012;491(7424):399–405.

[48] Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq 2008;5(7):1–8.

[49] Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, et al. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 2008;453(7199):1239–43. <http://www.ncbi.nlm.nih.gov/pubmed/18488015> [cited 20.08.13].

[50] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.

[51] Chen Z, Duan X. Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. Methods Mol Biol 2011;733:93–103.

[52] Benes V, Blake J, Doyle K. Ribo-Zero Gold Kit: improved RNA-seq results after removal of cytoplasmic and mitochondrial ribosomal RNA. Nat Methods 2011 Nov;8(11). http://dx.doi.org/10.1038/nmeth.f.352.

[53] Cui P, Lin Q, Ding F, Xin C, Gong W, Zhang L, et al. A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing. Genomics 2010;96(5):259–65. <http://www.ncbi.nlm.nih.gov/pubmed/20688152> [cited 04.10.13].

[54] Hornett EA, Wheat CW. Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. BMC Genomics 2012;13(1):361. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3469347&tool=pmcentrez&rendertype=abstract> [cited 26.09.13].

[55] Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol 2011;29(7):644–52. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3571712&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[56] Schulz MH, Zerbino DR, Vingron M, Birney E. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. Bioinformatics 2012;28(8):1086–92. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3324515&tool=pmcentrez&rendertype=abstract>.

[57] Robertson G, Schein J, Chiu R, Corbett R, Field M, Jackman SD, et al. De novo assembly and analysis of RNA-seq data. Nat Methods 2010;7(11):909–12. <http://www.ncbi.nlm.nih.gov/pubmed/20935650> [cited 06.08.13].

[58] Zhao Q-Y, Wang Y, Kong Y-M, Luo D, Li X, Hao P. Optimizing de novo transcriptome assembly from short-read RNA-seq data: a comparative study. BMC Bioinformatics 2011;12(Suppl. 14):S2. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3287467&tool=pmcentrez&rendertype=abstract>.

[59] Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-seq. Bioinformatics 2009;25(9):1105–11. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2672628&tool=pmcentrez&rendertype=abstract>.

[60] Au KF, Jiang H, Lin L, Xing Y, Wong WH. Detection of splice junctions from paired-end RNA-seq data by SpliceMap. Nucleic Acids Res 2010;38(14):4570–8. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2919714&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[61] Wang K, Singh D, Zeng Z, Coleman SJ, Huang Y, Savich GL, et al. MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. Nucleic Acids Res 2010;38(18):e178. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2952873&tool=pmcentrez&rendertype=abstract> [cited 07.08.13].

[62] Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 2010;26(7):873–81. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2844994&tool=pmcentrez&rendertype=abstract> [cited 07.08.13].

[63] De Bona F, Ossowski S, Schneeberger K, Rätsch G. Optimal spliced alignments of short sequence reads. Bioinformatics 2008;24(16):i 174–80. <http://www.ncbi.nlm.nih.gov/pubmed/18689821> [cited 14.08.13].

[64] Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 2010;28(5):511–5. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3146043&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[65] Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, et al. *Ab initio* reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol 2010;28(5):503–10. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2868100&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[66] Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 2008;18(9):1509–17. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2527709&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[67] Bullard JH, Purdom E, Hansen KD, Dudoit S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. BMC Bioinformatics 2010;11:94. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2838869&tool=pmcentrez&rendertype=abstract>.

[68] Dillies M-A, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, et al. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Brief Bioinform 2012. <http://www.ncbi.nlm.nih.gov/pubmed/22988256>.

[69] Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 2012;7(3):562–78. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3334321&tool=pmcentrez&rendertype=abstract> [cited 08.08.13].

[70] Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010;11(10):R106. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3218662&tool=pmcentrez&rendertype=abstract [cited 06.08.13].

[71] Robinson MD, McCarthy DJ, Smyth GK. EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26(1):139–40. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2796818&tool=pmcentrez&rendertype=abstract [cited 06.08.13].

[72] Hardcastle TJ, Kelly KA. BaySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics 2010;11:422. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2928208&tool=pmcentrez&rendertype=abstract.

[73] Barski A, Cuddapah S, Cui K, Roh T, Schones DE, Wang Z. Resource high-resolution profiling of histone methylations in the human genome. Cell 2007;129:823–37.

[74] Frommer M, Mcdonald LE, Millar DS, Collist CM, Watt F, Griggt GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci 1992;89(March):1827–31.

[75] Cheung M-S, Down TA, Latorre I, Ahringer J. Systematic bias in high-throughput sequencing data and its correction by BEADS. Nucleic Acids Res 2011;39(15):e103. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3159482&tool=pmcentrez&rendertype=abstract [cited 21.09.13].

[76] Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res 2012:1813–31.

[77] Fejes AP, Robertson G, Bilenky M, Varhol R, Bainbridge M, Jones SJM. FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. Bioinformatics 2008;24(15):1729–30. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2638869&tool=pmcentrez&rendertype=abstract> [cited 17.09.13].

[78] Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-seq (MACS). Genome Biol 2008;9(9):R137. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2592715&tool=pmcentrez&rendertype=abstract> [cited 20.09.13].

[79] Valouev A, Johnson DS, Sundquist A, Medina C, Anton E, Batzoglou S, et al. Genome-wide analysis of transcription factor binding sites based on ChIP-seq data. Nat Methods 2008;5(9):829–34.

[80] Ji H, Jiang H, Ma W, Wong WH. Using CisGenome to analyze ChIP-chip and ChIP-seq data. Curr Protoc Bioinformatics 2011;33:2.13.1–2.13.45.

[81] Boyle AP, Guinney J, Crawford GE, Furey TS. F-seq: a feature density estimator for high-throughput sequence tags. Bioinformatics 2008;24(21):2537–8. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2732284&tool=pmcentrez&rendertype=abstract> [cited 16.09.13].

[82] Jothi R, Cuddapah S, Barski A, Cui K, Zhao K. Genome-wide identification of *in vivo* protein-DNA binding sites from ChIP-seq data. Nucleic Acids Res 2008;36(16):5221–31. <http://www.pubmedcentral.

nih.gov/articlerender.fcgi?artid=2532738&tool=pmcentrez&rendertype=abstract> [cited 17.09.13].

[83] Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, et al. A CTCF-independent role for cohesin in tissue-specific transcription. Genome Res 2010;20(5):578–88. <http://genome.cshlp.org/content/20/5/578.abstract>.

[84] Rozowsky J, Euskirchen G, Auerbach RK, Zhang ZD, Gibson T, Bjornson R, et al. PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. Nat Biotechnol 2009;27(1):66–75. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2924752&tool=pmcentrez&rendertype=abstract> [cited 21.09.13].

[85] Micsinai M, Parisi F, Strino F, Asp P, Dynlacht BD, Kluger Y. Picking ChIP-seq peak detectors for analyzing chromatin modification experiments. Nucleic Acids Res 2012;40(9):e70. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3351193&tool=pmcentrez&rendertype=abstract> [cited 21.09.13].

[86] Xu H, Handoko L, Wei X, Ye C, Sheng J, Wei C-L, et al. A signal-noise model for significance analysis of ChIP-seq with negative control. Bioinformatics 2010;26(9):1199–204. <http://www.ncbi.nlm.nih.gov/pubmed/20371496> [cited 01.10.13].

[87] Xu H, Wei C-L, Lin F, Sung W-K. An HMM approach to genome-wide identification of differential histone modification sites from ChIP-seq data. Bioinformatics 2008;24(20):2344–9. <http://www.ncbi.nlm.nih.gov/pubmed/18667444> [cited 20.09.13].

[88] Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W. A clustering approach for identification of enriched domains from histone modification ChIP-seq data. Bioinformatics 2009;25(15):1952–8. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2732366&tool=pmcentrez&rendertype=abstract> [cited 21.09.13].

[89] Song Q, Smith AD. Identifying dispersed epigenomic domains from ChIP-seq data. Bioinformatics 2011;27(6):870–1. <http://bioinformatics.oxfordjournals.org/content/27/6/870.abstract>.

[90] Hower V, Evans S, Pachter L. Shape-based peak identification for ChIPsSeq. BMC Bioinformatics 2011;12(1):15. <http://www.biomedcentral.com/1471-2105/12/15>.

[91] Lan X, Bonneville R, Apostolos J, Wu W, Jin VX. W-ChIPeaks: a comprehensive web application tool for processing ChIP-chip and ChIP-seq data. Bioinformatics 2011;27(3):428–30. <http://bioinformatics.oxfordjournals.org/content/27/3/428.abstract>.

[92] Selker E, Tountas N, Cross S, Margolin B, Murphy J. The methylated component of the *Neurospora crassa* genome. Nature 2003;422:893–7.

[93] Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan S. Genome-wide-high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. Cell 2006;126:1189–201.

[94] Weber M, Hellmann I, Stadler M, Ramos L, Paabo S. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:457–66.

[95] Zilberman D, Gehring M, Tran R, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet 2007;39:61–9.

[96] Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics 2009;10:232. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2724425&tool=pmcentrez&rendertype=abstract> [cited 22.09.13].

[97] Harris EY, Ponts N, Levchuk A, Le Roch K, Lonardi S. BRAT: bisulfite-treated reads analysis tool. Bioinformatics 2010;26(4):572–3. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3716225&tool=pmcentrez&rendertype=abstract> [cited 23.09.13].

[98] Smith A, Chung W, Hodges E, Kendall J, Hannon G, Hicks J, et al. Updates to the RMAP short-read mapping software. Bioinformatics 2009;25:2841–2.

[99] Chen P, Cokus SJ, Pellegrini M. BS Seeker: precise mapping for bisulfite sequencing software. BMC Bioinformatics 2010;11:203.

[100] Lim J, Tennakoon C, Li G, Wong E, Ruan Y, Wei C. BatMeth: improved mapper for bisulfite sequencing reads on DNA methylation. Genome Biol 2012;13(10):R82. <http://genomebiology.com/2012/13/10/R82>.

[101] Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. Bioinformatics 2011;27(11):1571–2. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3102221&tool=pmcentrez&rendertype=abstract> [cited 22.09.13].

[102] Otto C, Stadler PF, Hoffmann S. Fast and sensitive mapping of bisulfite-treated sequencing data. Bioinformatics 2012;28(13):1698–704. <http://www.ncbi.nlm.nih.gov/pubmed/22581174>.

[103] Kreck B, Marnellos G, Richter J, Krueger F, Siebert R, Franke A. B-SOLANA: an approach for the analysis of two-base encoding bisulfite sequencing data. Bioinformatics 2012;28(3):428–9. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3268236&tool=pmcentrez&rendertype=abstract>.

[104] Krueger F, Kreck B, Franke A, Andrews SR. DNA methylome analysis using short bisulfite sequencing data. Nat Methods 2012;9(2):145–51. <http://www.ncbi.nlm.nih.gov/pubmed/22290186> [cited 23.09.13].

[105] Gruntman E, Qi Y, Slotkin RK, Roeder T, Martienssen RA, Sachidanandam R. Kismeth: analyzer of plant methylation states through bisulfite sequencing. BMC Bioinformatics 2008;9:371.

[106] Rohde C, Zhang Y, Reinhardt R, Jeltsch A. BISMA–fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. BMC Bioinformatics 2010; 11:230. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2877691&tool=pmcentrez&rendertype=abstract>.

[107] Kumaki Y, Oda M, Okano M. QUMA: quantification tool for methylation analysis. Nucleic Acids Res 2008;36(Web Server issue):W170–5. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2447804&tool=pmcentrez&rendertype=abstract> [cited 21.09.13].

[108] Benoukraf T, Wongphayak S, Hadi LHA, Wu M, Soong R. GBSA: a comprehensive software for analysing whole genome bisulfite sequencing data. Nucleic Acids Res 2013;41(4):e55. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3575794&tool=pmcentrez&rendertype=abstract> [cited 21.09.13].

[109] Berg R. The indigenous gastrointestinal microflora. Trends Microbiol 1996;4:430–5.

[110] Savage D. Microbial ecology of the gastrointestinal tract. Annu Rev Microbiol 1977;31:107–33.

[111] Kinross JM, Darzi AW, Nicholson JK. Gut microbiome–host interactions in health and disease. Genome Med 2011;3:14.

[112] Sartor RB. Microbial influences in inflammatory bowel diseases. Gastroenterology 2008;134:577–94.

[113] Aas J, Gessert CE, Bakken JS. Recurrent Clostridium difficile colitis: case series involving 18 patients treated with donor stool administered via a nasogastric tube. Clin Infect Dis 2003;36:580–5.

[114] Ley R. Obesity alters gut microbial ecology. Proc Natl Acad Sci 2005;102:11070–5.

[115] Ravel J. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci 2011;108(Suppl. 1):4680–7.

[116] Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone Ca, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci USA 2011;108(Suppl. 1):4516–22. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3063599&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[117] Claesson MJ, Wang Q, O'Sullivan O, Greene-Diniz R, Cole JR, Ross RP, et al. Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. Nucleic Acids Res 2010;38(22):e200. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3001100&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[118] Lazarevic V, Whiteson K, Huse S, Hernandez D, Farinelli L, Osterås M, et al. Metagenomic study of the oral microbiota by illumina high-throughput sequencing. J Microbiol Methods 2009;79(3):266–71. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3568755&tool=pmcentrez&rendertype=abstract> [cited 08.08.13].

[119] Baker GC, Smith JJ, Cowan DA. Review and re-analysis of domain-specific 16S primers. J Microbiol Methods 2003;55(3):541–55. <http://linkinghub.elsevier.com/retrieve/pii/S0167701203002276> [cited 16.09.13].

[120] Watanabe K, Kodama Y, Harayama S. Design and evaluation of PCR primers to amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. J. Microbiol. Methods 2001;44(3):253–62. <http://www.ncbi.nlm.nih.gov/pubmed/11240048>.

[121] Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 2007;73(16):5261–7. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1950982&tool=pmcentrez&rendertype=abstract> [cited 07.08.13].

[122] Garrity GM, Bell JA, Lilburn TG. Taxonomic outline of the prokaryotes. Bergey's Manual of Systematic Bacteriology. 2nd ed. New York: Springer-Verlag; 2004. http://dx.doi.org/10.1007/bergeysoutline200405 [Release 5.0].

[123] Quince C, Lanzén A, Curtis TP, Davenport RJ, Hall N, Head IM, et al. Accurate determination of microbial diversity from 454 pyrosequencing data. Nat Methods 2009;6(9):639–41<http://www.ncbi.nlm.nih.gov/pubmed/19668203> [cited 31.07.13].

[124] Burke C, Kjelleberg S, Thomas T. Selective extraction of bacterial DNA from the surfaces of macroalgae. Appl Env Microbiol 2009;75:252–6.

[125] Thomas T, Rusch D, DeMaere M, Yung P, Lewis M, Halpern A, et al. Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. ISME J 2010;4:1557–67.

[126] Chevreux B, Wetter T, Suhai S. Genome sequence assembly using trace signals and additional sequence information computer science and biology. Proc Ger Conf Bioinform 1999;99:45–56.

[127] Wooley JC, Godzik A, Friedberg I. A primer on metagenomics. PLoS Comput Biol 2010;6(2):e1000667. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2829047&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[128] Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. Microb Inform Exp 2012;2(1):3. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3351745&tool=pmcentrez&rendertype=abstract> [cited 06.08.13].

[129] Namiki T, Hachiya T, Tanaka H, Sakakibara Y. MetaVelvet: an extension of velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res 2012;40(20):e155. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3488206&tool=pmcentrez&rendertype=abstract> [cited 16.09.13].

[130] Peng Y, Leung HCM, Yiu SM, Chin FYL. Meta-IDBA: a de novo assembler for metagenomic data. Bioinformatics 2011;27(13):i94–i101. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3117360&tool=pmcentrez&rendertype=abstract> [cited 17.09.13].

[131] Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, et al. The metagenomics RAST server—a public resource for the automatic phylo- genetic and functional analysis of metagenomes. BMC Bioinformatics 2008;9:386.

[132] Wang J, Qi J, Zhao H, He S, Zhang Y, Wei S, et al. Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. Sci Rep 2013;3:1843. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3654486&tool=pmcentrez&rendertype=abstract> [cited 16.09.13].

[133] Shen P, Wang W, Chi A-K, Fan Y, Davis RW, Scharfe C. Multiplex target capture with double-stranded DNA probes. Genome Med 2013;5(5):50. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3706973&tool=pmcentrez&rendertype=abstract> [cited 29.09.13].

[134] Mertes F, Elsharawy A, Sauer S, van Helvoort JMLM, van der Zaag PJ, Franke A, et al. Targeted enrichment of genomic DNA regions for next-generation sequencing. Brief Funct Genomics 2011;10(6):374–86. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3245553&tool=pmcentrez&rendertype=abstract> [cited 19.09.13].

[135] Xu W, Seok J, Mindrinos MN, Schweitzer AC, Jiang H, Wilhelmy J, et al. Human transcriptome array for high-throughput clinical studies. PNAS 2011;108:3701–12.

This page is intentionally left blank

Chapter 9

Proteomics

Holger Husi[a] and Amaya Albalat[b]
[a]BHF Glasgow Cardiovascular Research Centre, ICAMS University of Glasgow, UK,
[b]School of Natural Sciences University of Stirling, UK

Chapter Outline

9.1 Overview 147
9.2 Data Acquisition 148
9.2.1 Sample Types and Sample Preparation Requirements/Protocols 148
9.2.2 Analytical Platforms: Separation and Detection Techniques 152
9.2.3 Ionization Techniques 157
9.2.4 Identification of Peptides or Sequencing MS/MS 158
9.2.5 Quantitative Approaches 159
9.3 Data Analysis 163
9.3.1 Downstream Processing 163
9.3.2 Biomarker Pattern Analysis 164
9.3.3 Data Integration and OMICS Analysis 167
9.4 Discovery and Validation 167
9.4.1 Cancer Markers 167
9.4.2 Cardiovascular Diseases 167
9.4.3 Kidney Diseases 170
9.4.4 Preeclampsia 171
9.4.5 Neurological Complexes 171
9.5 Personalized Medicine 172
9.6 Conclusions 172
Acronyms 173
References 173

9.1 OVERVIEW

Differences between physiological states at the cellular level are reflected by the expression level and state of all proteins. Therefore, proteomics, which aims at this, can in principle provide the most relevant single dataset that characterizes a biological system [1]. Mainly for this reason, proteomics has had a tremendous impact in biological research over the last 15 years; specifically, cancer research has benefited from this biological research discipline [2]. Figure 9.1 is a “word cloud” generated by analysis of website HTML tags associated with proteomics; and it demonstrates where proteomic discoveries have been most employed to date—namely, cancer research, drug and drug target discovery, and biomarker research.

Initially, study of the proteome was attempted using mainly two-dimensional gel electrophoresis. However, nowadays most proteomic studies use mass spectrometry (MS). Using MS, and very small amounts of biological samples or biofluids which are highly complex, it is possible to quickly identify peptides by matching their fragmentation profiles with available sequence databases. This approach is based on extensive use of tandem MS. In this case, full-scan MS is carried out on intact molecular ions (proteins or peptides) that are then fragmented successively to produce fragmentation scans, which are the key to obtaining primary sequence information (MS^n scans) [3].

A typical proteomic workflow starts with the solubilization of proteins from a sample that can then be digested enzymatically or chemically into peptides (bottom-up analysis) or analyzed directly as intact proteins (top-down analysis). Normally, a bottom-up approach is used for the study of complex samples (whole tissues, biofluids, etc.) in a high-throughput manner. In this bottom-up approach, also known as “shotgun proteomics,” the identification of proteins is achieved by utilizing the combination of tandem MS (or MS/MS data) and database searching [4]. In this context, we can find programs such as SEQUEST, where experimental MS/MS data are compared to the predicted in silico-generated peptide fragmentation patterns.

In addition to SEQUEST, many other methods have been developed to improve this process, as several challenges associated with bottom-up proteomics are still to be fully resolved, including probabilistic scoring schemes [5], implementing additional search criteria [6], and the use of identified spectra to bootstrap the database search [7].

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00009-8
© 2014 Elsevier Inc. All rights reserved.

discovery
Mass
Research
spectrometry
Functional
Drug
Protein
Cancer
Imaging
Molecular
Biomarker
Chromatography
expression
throughput
High
Database
Data
Proteomics
proteomics

FIGURE 9.1 **Proteomics word cloud.** Tag words associated with "Proteomics" from websites found on the www are depicted according to their frequencies attached to webpage-associated information to calculate the font-sizes in the figure.

Currently, there is no public repository of proteomic data. However, it is expected that public repositories will be available in the future that will promote and reinforce the need for standardization of data formats, as well as increase the availability of proteomic data to other scientific disciplines.

The main advantages of the bottom-up approach include a higher resolution of peptides compared to proteins in the separation step as well as a high sensitivity, which is of great importance because proteins, in contrast to nucleotides, cannot be amplified. Furthermore, with the bottom-up approach it is possible to chemically modify peptides to improve their quantification. ICAT [8], ^{18}O labeling [9], and tandem mass tags [10] are suitable techniques for peptide-based methodologies such as MudPIT. However, the bottom-up approach also has certain limitations that should be taken into consideration, such as a loss of information because of limited protein sequence coverage by the identified peptides and the loss of labile post-translation modifications, which can be very important in biomarker research. Furthermore, there can be ambiguity in the identification of isoforms or conserved domains, which are sometimes found in many proteins because of the multiple origins of identified peptides.

Top-down methods are based on the study of intact protein masses and their fragments for identification. They have progressed mainly because of the development of fragmentation reactions, including electron capture dissociation (ECD) and electron transfer dissociation (ETD), which can yield a more complete backbone protein fragmentation with more sequencing information while retaining important labile post-translational modifications (PTMs) [4]. In this approach, proteins are not chemically or enzymatically digested and therefore it is necessary to analyze simpler, or "cleaner," samples for their successful analysis. Data analysis utilizes the expressed sequence tag (EST) or de novo methods [11].

The main advantages of the top-down approach are higher sequence coverage of target proteins and more accurate characterization of PTMs [12]. The more extensive sequence coverage of the top-down approach reduces problems associated with peptide-to-protein mapping and improves isoform identification, which is an issue for bottom-up approaches [13]. Top-down approaches also have more reliable protein quantification [14], where intact protein abundances are measured directly. However, there are some technical limitations of the top-down approach, and therefore both top-down and bottom-up approaches should be perceived as complementary. In particular, the separation of intact proteins before MS analysis is very challenging and rather limited, although recent studies [15,16] have used the top-down approach for complex mixture analysis. Furthermore, higher mass accuracy instruments are needed for analysis (FTMS and LTQ-Orbitrap).

Figure 9.2 summarizes the most common processes and workflows used in published proteomics studies. In this chapter, we cover the main methodologies used for sample preparation, separation of proteins using gel and non-gel-based methods and possible methods for the identification and quantification of proteins/peptides. Finally, considerations for data analysis and evaluation are discussed.

9.2 DATA ACQUISITION

9.2.1 Sample Types and Sample Preparation Requirements/Protocols

In a research environment, the collection of biological samples (tissues or biofluids) for proteomic studies can be less problematic (more controlled parameters and less time) compared to large clinical trials. However, with the increasing evidence that pharmacogenomics, biomarker discovery, and proteomics play an important role in disease characterization at different levels—from epidemiology and diagnostic and prognostic classification to response prediction—therapy-appropriate samples for molecular research have become essential [17]. This increasing demand for personalized medicine is now being recognized not only by the pharmaceutical industry, but also by regulatory agencies and the general public. Such approaches are highly dependent on accumulating enough biological specimens to validate any findings.

Internationally, it can be predicted that biological sample collection will become increasingly accepted based on an ever growing body of molecular information that aids in understanding the health benefits of undertaking

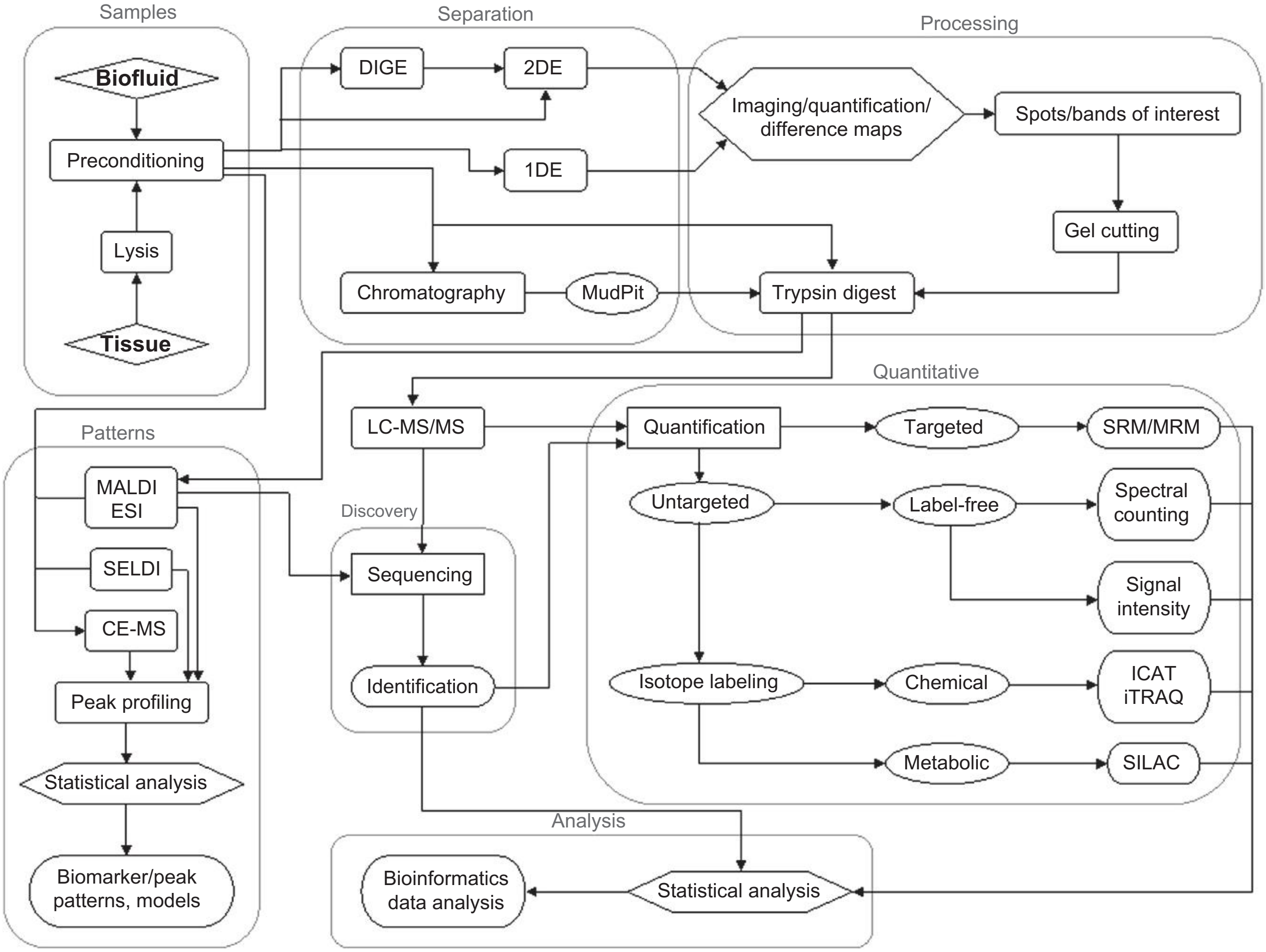

Samples
Biofluid
Preconditioning
Lysis
Tissue
Separation
DIGE
2DE
1DE
Chromatography
MudPit
Processing
Imaging/quantification/ difference maps
Spots/bands of interest
Gel cutting
Trypsin digest
Patterns
MALDI ESI
SELDI
CE-MS
Peak profiling
Statistical analysis
Biomarker/peak patterns, models
LC-MS/MS
Discovery
Sequencing
Identification
Quantitative
Quantification
Targeted
SRM/MRM
Untargeted
Label-free
Spectral counting
Signal intensity
Isotope labeling
Chemical
ICAT iTRAQ
Metabolic
SILAC
Analysis
Bioinformatics data analysis
Statistical analysis

FIGURE 9.2 **Proteomics workflow.** The most common processing flows and workflows are indicated; however, many more combinations and alternatives are possible.

such studies. However, for some areas of research, such as genome profiling, the opposite might hold true because of regional legislation and reservations. Globally accepted standardization of appropriate procedures to routinely collect samples has helped in acceptance of pharmacogenomic applications in clinical trials [18]. The type of sample to be analyzed often directly influences which methods can be employed in proteomic downstream analysis. Table 9.1 lists a few advantages and disadvantages of the main sample types used in proteomics as well as selected generic protocols for sample handling and conditioning for MS-based analysis.

9.2.1.1 *Urine*

Urine is an ideal source for biomarker research aimed at clinical application because it can be easily collected in large quantities using noninvasive methods. Proteins in the urine are mainly composed of plasma proteins that have passed through the glomerular filtration barrier and proteins secreted from the kidney and urinary tract [19]. A protocol for urine collection has been published following the recommendations of the Human Kidney and Urine Proteome Project (HKUPP) and the European Urine and Kidney Proteomics, EuroKUP initiative (for further and more detailed information, see www.eurokup.org/sites/default/files/StandardProtocolforUrineCollection.pdf). Ideally, the sample is mid-stream of the second morning urine or morning random-catch urine. It is also recommended that urine samples be collected in sterile or clean urine collectors. The following protocol has been adapted from the literature [20].

1. Collect at least 3 ml of midstream urine per sample.
2. Immediately spin at 1000 × g for 5 min at room temperature.
3. Aliquot the supernatant into 1, 10, or 50 ml fractions (depending on downstream application) on dry ice.
4. Store aliquots at −80 °C ideally, or store them at −20 °C. Time until freezing should be limited to 3 h.

TABLE 9.1 Sample Source Advantages and Limitations

Sample Source	Advantages	Limitations
Urine	Easy and noninvasive collection, obtained in large volume	Fluctuation of protein concentration over time and according to renal integrity; presence of salts and proteins in low concentration
Plasma	Larger final volume of fluid after processing than with serum; more rapidly processed than serum; more stable than serum because of inhibition of coagulation cascade	Sample dilution in citrate tube; possible interference of EDTA with detection method
Serum	Established sample banks often composed of serum aliquots (retrospective studies), proteins, and peptides that "survive" to the clotting procedure exhibit stability that can be exploited in routine clinical applications	Presence of various products derived from coagulation cascade; biomarker with poor stability during coagulation process will not be detected in serum; possible influence of disease on coagulation process
Saliva	Easy and noninvasive sampling	Low volume collected; presence of many proteases and unspecific materials such as food residues or microorganisms; level of certain plasma proteins not reflected
Tissue	Whole biological systems analysis possible; scaleable	Sampling and quantities obtainable can be a problem; identification of interesting features can be hidden beneath high-abundance molecules

5. Thaw samples at 4 °C. If thawing and refreezing occur, always keep a record of this event. Avoid freeze-thaw cycles if possible or limit them to two.
6. Centrifuge at 1000 × g and 3000 × g for 5–10 min to remove precipitated material.
7. If necessary, measure total protein concentration or creatinine level and standardize by dilution in appropriate buffer.

Thereafter, preprocessing or solubilization of urine proteins ultimately depends upon the method of detection. Regardless of the method used, a clean-up or purification step is recommended to reduce the concentration of salt or other interfering compounds such as urea. Moreover, the depletion of human serum albumin can help to reduce the complexity of urine and improve the sensitivity of low-abundance proteins [21]. Methods for preprocessing urine samples for gel-based applications have been published [21], and there are methods available for non-gel applications (MALDI, LC-MS, and CE-MS) that readers are encouraged to review [20,22].

9.2.1.2 *Plasma*

Blood in general, and plasma in particular, has been the biofluid of choice for many proteomic studies, especially those aimed at biomarker research. The undeniable importance of blood is due to its function as a carrier for all molecules to and from all tissues and organs in the body. Plasma is collected in a tube containing anticoagulant—commonly EDTA, although citrate and heparin have also been used. It contains not only the traditional "plasma proteins," but also tissue proteins and numerous distinct immunoglobulin sequences [23]. However, because of the amount of information contained in this biofluid and a wide dynamic range of the present molecules (more than ten orders of magnitude in concentration) [23], analysis of all compounds in plasma is not possible with available technologies. For this reason, compound enrichment and/or depletion procedures are needed as sample preprocessing steps.

Therefore, preprocessing methods, such as centrifugal ultrafiltration, solid-phase extraction with its different formats, and organic solvent extraction, among others, are possibilities that readers are encouraged to review and test prior to a final experimental design [24]. Standardization of the preanalytical phase, including sample collection, sample preparation, and sample storage, is of paramount importance in all biofluids but especially in plasma, where differences in proteolytic activities during the preanalytical phase can have a substantial influence on the proteomic final results, possibly leading to inappropriate conclusions [25]. The following is an adapted protocol [26,27].

1. Collect blood samples in EDTA tubes.
2. Incubate vertically at 4 °C for 30 min and centrifuge at 3000 × g for 20 min at 4 °C.
3. Store plasma sample supernatants at −80 °C for further analysis.
4. The following steps (5–12) remove high-abundance proteins from plasma using the Multiple Affinity Removal column (10 × 100 mm column) and proprietary buffer system (MARS^{TM}, Agilent).
5. Equilibrate the column for 4 min at room temperature with a flow rate of 1 ml/min.
6. Dilute 200 μl of plasma with 600 μl of HPLC MARS buffer A.
7. Centrifuge diluted plasma through a Spin-X centrifuge 0.45 μm filter, 1 min at 16,000 × g to remove particulates.
8. Inject the diluted sample at a flow rate of 0.5–1 ml/min.
9. Collect the flowthrough fraction in the 10.5–14.5 min retention time range. This corresponds to the low-abundance protein fraction.
10. Dialyze the flowthrough fraction with ammonium bicarbonate and lyophilize. Store at −80 °C for later analysis.
11. Elute bound proteins by adding HPLC MARS buffer B at a flow rate of 3 ml/min for 7.5 min.
12. Resuspend lyophilized fractions in TEAB buffer and determine protein concentration.

9.2.1.3 Serum

While plasma is collected in a tube containing anticoagulant, serum is collected in a tube containing no additives (in some cases, a form of micronized silica can be used to promote clotting) [28]. The coagulation step has an effect on the composition of the proteins present in serum compared to plasma, which should be taken into consideration. The following is an adapted method [29].

1. Collect blood in a serum separator tube and allow it to clot for 30 min at room temperature.
2. Centrifuge for 10 min at 1000 × g.
3. Aliquot the supernatant serum before flash-freezing at –80 °C.
4. Thaw the serum aliquot by submersion in a water bath at room temperature and process as soon as possible. Note that the quality of the sample deteriorates with each freeze-thaw cycle.
5. Remove particulates by centrifugation at 10,000 × g for 10–15 s.

Regarding sample pretreatment, the challenges associated with serum are similar to those with plasma. Therefore, it is necessary to deplete highly abundant proteins and/or isolate and enrich low-molecular-weight proteins in order to detect candidate biomarkers and to avoid a loss of assay robustness and sensitivity.

9.2.1.4 Saliva

Saliva is an attractive biofluid for proteomic studies because of the abundance of proteins it contains, which can be very easily obtained inexpensively, safely, and noninvasively [30]. Many proteins contained in saliva have been shown to be of clinical importance [31]. As with other biofluids, standardization of saliva sample collection and pretreatment are needed to obtain reliable and significant results. The following is an adapted protocol [32].

1. Aliquot saliva samples into 330 μl fractions on ice.
2. Add protease inhibitors (aprotinin, vanadate, and PMSF or other cocktail) if required and store at –80 °C.
3. Omit step 2 if supernatant saliva needs to be generated. This is done by centrifugation at 2600 × g for 15 min at 4 °C, addition of protease inhibitors to the supernatant only if required, and storage of all fractions and the pellet at –80 °C.

9.2.1.5 Tissue

The proteomic study of tissue is a direct and very informative approach that is particularly useful for studying the pathophysiological events that take place in a diseased condition. From a biomarker perspective, the main limitation of tissue is that it is not generally accessible, especially from control or healthy organisms for monitoring purposes. In this chapter, we present a commonly used method for the bottom-up study of the tissue proteome. The main advantages of this protocol are that the protein digests that it obtains are free of nucleic acids and other cell components; it can be applied to samples containing high concentrations of detergent; and the sample concentration is kept high as there is no precipitation. This protocol is adapted from the filter-aided sample preparation (FASP) method [33].

1. Lyse cells or tissues in lysis buffer (i.e., 4% SDS and 0.1 M DTT, or 2% SDS and 0.1 M DTT, respectively, in 0.1 M Tris-HCl, pH 7.6 buffer) using a 1:10 sample to buffer ratio by sonication or blending; incubate at 95 °C for 3–5 min.
2. Sonicate to reduce the viscosity of the sample and disrupt DNA aggregates.
3. Centrifuge at 16,000 × g for 5 min to pellet unbroken cells and other aggregates.
4. Mix 1 mg of protein (approximately 30 μl) with 200 μl of 8 M urea in 0.1 M Tris pH 8.5 and concentrate in an Amicon (Millipore), Hydrosart (Sartorius), or Vivaspin (Sartorius) centrifugal concentrator at 14,000 × g for 40 min.
5. Repeat the dilution and concentration step with the retenate; discard the flowthrough.

6. Add 100 µl of 0.05 M iodoacetamide in 8 M urea and 0.1 M Tris pH 8.5; mix for 1 min at 600 rpm and incubate without mixing for 20 min.
7. Centrifuge filter unit at 14,000 × g for 10 min.
8. Add 100 µl of 8 M urea in 0.1 M Tris pH 7.9, and concentrate by centrifugation at 14,000 × g for 15 min. Repeat this step twice.
9. Add 100 µl of 50 mM ammonium bicarbonate to the filter unit and centrifuge at 14,000 × g for 10 min. Repeat this step twice.
10. Digest the proteins by adding 40 µl of 50 mM ammonium bicarbonate containing trypsin, with a recommended ratio of enzyme to protein of 1:100; mix at 600 rpm for 1 min. The peptide content can be estimated by UV light spectral density at 280 nm using an extinction coefficient of 1.1 of 0.1% ($g\,l^{-1}$) solution.
11. Incubate the unit in a wet chamber overnight.
12. Transfer the filter unit to new collection tubes.
13. Centrifuge the filter unit at 14,000 × g for 10 min.
14. Add 40 µl of 50 mM ammonium bicarbonate to the filter units and centrifuge at 14,000 × g for 10 min.
15. Acidify the sample with trifluoroacetic acid or formic acid and desalt the filtrate.
16. Desalt small amounts of digest for direct LC-MS analysis on StageTips or ZipTips.
17. Desalt large amounts of peptide mixtures for off-gel separation on SPE cartridges according to steps 18–25.
18. Place a 3 ml MILI-SPE extraction disk cartridge (C18-SD) in a 15 ml conical tube.
19. Add 1 ml of methanol and centrifuge at 1500 × g for 1 min.
20. Add 0.5 ml of 0.1% trifluoroacetic acid, 70% acetonitrile in water; centrifuge at 1500 × g for 1 min.
21. Add 0.5 ml of 0.1% trifluoroacetic acid in water and centrifuge at 1500 × g for 1 min.
22. Load the filtrate and centrifuge at 150 × g for 3 min.
23. Add 0.5 ml of 0.1% trifluoroacetic acid in water and centrifuge at 150 × g for 3 min.
24. Transfer the cartridge to a new tube, add 0.5 ml 70% acetonitrile in water, and centrifuge at 150 × g for 3 min.
25. The eluate contains desalted peptides.

9.2.2 Analytical Platforms: Separation and Detection Techniques

Proteomic workflows based on MS analysis are ideally suited to investigate large and complex protein networks and dynamic molecular interactions that modulate molecular, cellular, and physiological processes in an unbiased and systematic fashion [34]. However, more often than not methodologies prior to MS analysis are used—for example, to reduce the complexity of biological systems, or to render samples compatible with MS analysis. Polyacrylamide gel electrophoresis (PAGE), capillary electrophoresis (CE), and liquid chromatography (LC) are such strategies routinely used in the separation and analysis of proteins and peptides [35].

All of these methodologies have led to further development of instrumentations, novel methods in data acquisition, and bioinformatics data analysis software tools [4,36]. Individually, both gel-based and gel-free methods have unique properties that can complement each other. The modular setup of many of these separation and MS analysis tools allows various adaptations in a multitude of ways [37], and combinations can result in advantages in analysis of complex biological samples. This has been demonstrated in numerous studies [38]. Table 9.2 lists the most common platforms used in proteomic analysis.

9.2.2.1 Gel-Based Approaches

Separation of polypeptides is often needed prior to mass spectrometric analysis to reduce the complexity of or to establish a focus for the molecules of interest. A traditional approach in most biology labs involves the use of gel-based systems, using either one-dimensional or two-dimensional separations.

One-Dimensional

One of the most common laboratory practices involves the separation of proteins by their migratory behavior through a molecular mesh by applying an electric current. This is known as one-dimensional gel electrophoresis or 1D gel analysis; it was first introduced by Laemmli in 1970 [39]. This method gives information about the purity of protein samples, yield of fractionation techniques, and molecular size and relative quantity of individual proteins. Subsequent recovery of separated molecules enables further analysis, such as amino acid sequence determination and ultimately protein identification by MS [40].

A variety of gel formats are available. Large gels are mainly used to separate complex mixtures or for specific applications such as 2D-PAGE. Small gels, called minigels, are the mainstay of most proteomic applications. The benefit of small gels is that they are relatively quick to perform, easy to use, relatively cheap to run, and, most importantly, commercially available in precast formats. The modified protocol, shown next, which is mainly used for the separation of small proteins and peptides of less than 20 kDa, includes the necessary chemical modifications of proteins/peptides required for subsequent MS analysis [41], and involves the use of commercially available precast gels.

1. Add 15 µl Tris-Tricine loading buffer (2× concentrated, BioRad) to 15 µl of the sample.
2. Reduce by adding DTT (1 mM final concentration) and incubate for 10 min at 70 °C. Alkylate by

TABLE 9.2 Comparison of Platforms in Proteomics Research

Platform	Advantages	Limitations
1D gel-based	Inexpensive; simultaneous separation of many proteins; large mass range	Time-consuming; not suitable for analyzing smaller peptides, hydrophobic, very large or very small proteins
2D gel-based	Available in most protein labs; simple to perform	Time-consuming; no automation; problematic with hydrophobic proteins; mass range >10 kDa
DIGE	Higher reproducibility and quantification than 2DE	Expensive, same limitations as 2DE; requires specific imaging system
CE-MS	Automated; high sensitivity; multidimensional; low sample volumes required; reproducible; biomarker identification	Dynamic range restricted to proteins <20 kDa; protein identification not possible
LC-MS	Automated; multidimensional; high sensitivity	Quantitative analysis not simple; sensitive to interfering compounds; measurements of post-translational modifications requires additional tools
MudPIT	Automated; multidimensional; high sensitivity; detection of low-abundance proteins	Quantitative analysis not simple; sensitive to interfering compounds; high cost; additional steps
SELDI-MS	Simple to use; high-throughput, automated; low sample volumes required; enrichment of specific low-abundance molecules; biomarker identification; high dynamic range	Low–medium-resolution MS; limited information content; reproducibility and accuracy relatively low on older instruments; no protein identification; only relative quantification
ESI	Small sample volume; widely available	Very sensitive to interfering compounds
MALDI	High-throughput; high sensitivity and reproducibility	Machine-dependent; dynamic-range limitations; protein identification not easy
iTRAQ	Relative quantification of low-abundance molecules and hydrophobic samples; good reproducibility	Expensive; time-consuming; not easy to use
SILAC	Relative quantification, proteomics, and metabolomics; low-abundance molecules	Expensive; time-consuming; not easy to use; limited applicability
Label-free, SRM, MRM	Relative and absolute quantification; reproducibility	Expensive; software-dependent analysis; uncertain accuracy

adding 2 mM iodoacetamide (final concentration) and incubate 30 min in the dark. Quench by adding DTT to a final concentration of 50 mM. Heat samples at 95 °C for 5 min and load on 16% Tris-Tricine peptide or Tris-Glycine gel; run 2 h at 100 V (large gels) or 45 min at 200 V for minigels, or until the loading dye reaches the end of the gel.

3. Fix the gel by incubating for 30 min in 40% methanol, 10% glacial acetic acid.
4. Rinse gel with water.
5. Incubate in colloidal Coomassie for 2 h, gently shaking.
6. Destain 4 × 50 ml water 30 min each.
7. Excise bands on a light table and store at −40 °C.
8. Destain completely by first adding 50 µl 50 mM ammonium bicarbonate to gel bands and incubate for 30 min at 37 °C.
9. Add 100 µl acetonitrile and incubate for 30 min at 37 °C.
10. Remove liquid and repeat until the gel bands are completely destained.
11. Add 50 µl acetonitrile and incubate 10 min at room temperature to shrink the gel pieces; remove the liquid and air-dry the gel pieces for 10 min. Store at −40 °C if required.
12. Digest: reconstitute a vial of Trypsin (25 µg) in 250 µl 0.1% trifluoroacetic acid, dilute 15 µl of this stock in 100 µl cold ammonium bicarbonate buffer, add 5–10 µl to gel (10 for visible bands, 5 for clear regions) and add 50–100 µl ammonium bicarbonate to cover the gel pieces.
13. Incubate for 1 h at 37 °C, then overnight (up to 18 h) at 25 °C.

14. Add 100 µl acetonitrile and incubate for 30 min at 37 °C.
15. Spin and retain the supernatant in a new vial.
16. Add 50 µl 0.5% formic acid to the gel and incubate for 30 min at 37 °C; then add 100 µl acetonitrile and incubate for 30 min at 37 °C and combine supernatant with previous. Repeat this step.
17. Lyophilize the combined supernatant and store at −40 °C.

Two-dimensional and 2D-PAGE

Two-dimensional polyacrylamide electrophoresis (2D-PAGE), introduced in 1975 by O'Farrell [42], combines both isoelectric focusing (IEF) and 1D SDS-PAGE analysis. The first dimension separates biomolecules based on their isoelectric point (pI); the second dimension based on their molecular sizes [42]. This technique can, in theory, resolve any proteins present in a sample, including those containing PTMs or transiently expressed molecules [43]. Several disadvantages limit the general use of this methodology, such as dynamic range constraints, difficult to resolve proteins containing large hydrophobic regions, and often poor gel-to-gel reproducibility. Additionally, though high-resolution 2D-PAGE is capable of resolving 2000–5000 protein spots per gel [44], this is far below the expected several million potential isoforms in a complex biological system. The following generic protocol is a modified version of commercial IEF gel (IPG), 2D-PAGE kit (GE Healthcare), [45] and colloidal Coomassie staining [46].

1. Rehydrate a 24 cm IPG strip pH 3–11 nonlinear gradient with 450 µl of rehydration buffer containing 97 mM DeStreak (GE Healthcare).
2. Dilute the sample in a rehydration buffer containing 50 mM DTT and load onto the previously rehydrated IPG strip via cup loading.
3. Run the first dimension: the IEF can be carried out on an IPGphor II IEF system (GE Healthcare) up to a total of 42 kVh.
4. Second dimension preconditioning: equilibrate the strip in 10 ml of reducing SDS equilibration buffer 15 min at room temperature with agitation. Then eliminate the reducing solution and incubate with alkylating SDS equilibration buffer for an additional 15 min.
5. Second dimension: carry out on a 12% polyacrylamide gel at 2 W/gel (for overnight running) or after sample entry in the running gel (40 min); increase to 20 W/gel (max. 100 W) for in-day gel running, until the loading dye reaches the end of the gel.
6. Fix the gel for 1 h in an aqueous solution containing 2% (v/v) glacial acetic acid and 40% (v/v) methanol.
7. Stain for 1 h or overnight in 125 ml of a solution containing 20 ml colloidal Coomassie stock, 80 ml water and 25 ml methanol.
8. Rinse gel in an aqueous solution containing 5% (v/v) glacial acetic acid and 25% (v/v) methanol.
9. Destain for up to 1 h with 25% (v/v) methanol.
10. Cut spots of interest, digest according to the previous protocol, and analyze by MS and MS/MS.

DIGE

The difference gel electrophoresis (DIGE) method was developed to overcome some of the problems associated with 2D-PAGE, such as analytical variations, and to address the quantitative aspect of differential-expression experiments. This powerful quantitative technology has proven to be an important tool to assess and simultaneously measure with statistical confidence individual protein abundance changes for hundreds, if not thousands of intact molecules, in replicates over multiple variables. The technique uses up to three spectrally distinguishable fluorescent dyes (CyDyes Cy2, Cy3, and Cy5) in a multiplexing analytical run, where individual samples are labeled with one of the dyes, then combined, and then co-resolved in the same gel. Such a multiplexing step is beyond the capabilities of standard 2D-PAGE and allows for a direct quantitative comparison between samples analyzed on the same gel [47]. The following method has been adapted from the literature [48] using a commercial 2D analysis, a DIGE kit, and a software analysis solution (GE Healthcare). It includes further steps for alternative staining [49] and digestion [50] protocols.

1. Isolate proteins from samples, precipitate using the 2D clean-up kit, and resuspend in lysis buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 30 mM Tris-base; measure protein concentration.
2. Prepare CyDye reagents by adding N,N-dimethyl formamide (DMF) to each of three CyDye vials to obtain a final working concentration of 100 pmol/µl.
3. Spin down samples and prepare individual samples with 50 µg of protein extract.
4. Label individual samples with Cy3 or Cy5 dye by adding 2–4 µl (200–400 pmol) of either Cy3 or Cy5 working dilution for 30 min on ice in the dark.
5. Quench the reaction with 2 µl of 10 mM lysine for 10 min on ice in the dark.
6. For the internal standard, prepare a pool with 25 µg of protein extract from each individual sample. Label with Cy2 working dye on ice for 30 min in the dark and quench the reaction with 6 µl of 10 mM lysine for 10 min.
7. For each gel, combine the quenched Cy3- and Cy5-labeled samples and add 1/3 of the quenched Cy2-labeled pooled mixture.
8. Perform IEF and SDS-PAGE separation based on the protocol for 2D-PAGE, steps 1–5.
9. Image acquisition: for each gel, obtain the Cy2-, Cy3-, and Cy5-labeled channel image at excitation/emission

wavelength values of 488/520, 523/580, and 633/670 nm, respectively.

10. Analyze the images using DeCyder software (GE Healthcare) for spot detection and quantification and intergel matching and statistics: (a) DIA (differential in-gel analysis) for spot detection, which compares all three fluorescence images obtained from each gel and performs co-detection, background removal, and normalization; (b) BVA (biological variation analysis) for matching multiple 2D-DIGE gels, which provides statistical analysis of protein abundance changes; (c) EDA (extended data analysis) for advanced statistical analysis to generate the final list of candidate spots, PCAs (principal component analyses), and hierarchical clustering.
11. If MS analysis is to be performed, rerun samples without CyDyes (2D-PAGE, steps1–5) and stain gels using silver staining.
12. For silver staining, fix the gels in aqueous 50% (v/v) methanol, 5% acetic acid, for 20 min.
13. Wash gels in 50% (v/v) methanol for 10 min, then in water for 10 min.
14. Reduce for 1 min in 0.02% sodium thiosulfate (50 mg/250 ml).
15. Wash for 1 min in water twice.
16. Incubate with 0.1% silver nitrate (cold, 4 °C) (250 mg/250 ml) for 20 min.
17. Wash for 1 min in water twice.
18. Develop with 0.04% formalin in 2% anhydrous sodium carbonate (10 g/500 ml) in water.
19. Stop the reaction with 5% acetic acid.
20. Spot picking: match silver-stained gels against DIGE images and excise selected spots.
21. Reduce gel spots with 10 mM DTT in 50 mM ammonium bicarbonate.
22. Alkylate with 55 mM iodoacetamide in 50 mM ammonium bicarbonate.
23. Rinse gel pieces with 50 mM ammonium bicarbonate and acetonitrile and dry.
24. Digest proteins with trypsin at a final concentration of 7.5 ng/μl in 50 mM ammonium bicarbonate for 10 h at 37 °C.
25. Lyophilize supernatant and keep at 4 °C for later MS analysis.

9.2.2.2 *Non-Gel-Based Approaches*

Many biological samples contain a complex mixture of protein isoforms with or without PTMs, cleavages, and/or splice variants, as well as abundances that span a dynamic range of at least 12 orders of magnitude [51]. As a result, decomplexing of biological samples becomes a necessity for many applications, such as biomarker discovery. Gel separation, in particular 2D-PAGE, has, in principle, the capability but not the capacity to spatially resolve complex proteinaceous mixtures. Intrinsic limitations include problematic resolution of very small or very large proteins or physicochemical properties that are difficult to handle by gel analysis, as well as differences in abundance that can potentially mask proteins present in low quantities and also hinder the analysis necessary for biomarker research or functional characterization. Additionally, identification of proteins by in-gel digestion and extraction of resulting peptides or recovery of whole proteins after gel separation is far from straightforward and is often associated with considerable losses. As a consequence, alternative methods based on LC fractionation and separation techniques are increasingly employed [52].

CE-MS

Capillary electrophoresis (CE) is an analytical separation technique first mentioned in the scientific literature in the early 1980s [53]. CE separation is driven by an electric field applied to an open capillary tube; it can yield the rapid partitioning of hundreds of different molecules, which migrate at varying speeds through the system depending on charge and size [54]. The solutes can be observed by a detector as peaks, and quantitative information can be recorded, since the area of a peak is directly proportional to the concentration of the molecule. CE-MS, or the coupling of an MS instrument to CE, enables accurate determination of the mass of each peak in addition to the quantitative data. The coupling of CE and MS was first reported in the late 1980s [55] and has been shown to be an invaluable tool in biomarker discovery [56].

The interface of CE with an MS instrument can be done in a sheathless or sheath-flow setting online. Separated molecules are normally ionized after CE separation using electrospray ionization (ESI). The main benefits of this technique are that it can be used in fully automatic mode without compromising sensitivity and selectivity. The following generic protocol using a sheath-flow interface for the analysis of proteins/peptides with a molecular mass lower than 20 kDa was adapted from the literature [57]. For the analysis of intact proteins using CE-MS, readers are encouraged to look at alternative protocols [58].

1. Install the capillary tube (90 cm × 50 μm, uncoated fused silica capillary) in the CE instrument offline from the MS detector.
2. Condition the capillary with sodium hydroxide (1 M) at 50 psi pressure for 10 min, and then rinse the capillary with ammonium hydroxide (add 3.76 ml of ammonium solution at 25% and adjust the volume to 50 ml with deionized water), also applied at 50 psi for 10 min. Finally, rinse the capillary with running buffer at the same pressure for at least 20 min. After this conditioning step, connect the capillary online to the MS instrument.

3. If the CE is from Beckman, it must be run in "reverse mode" when connected to the MS system; therefore, an external detector adapter is necessary.
4. Before the injection of a sample, rinse the capillary with running buffer (add 10 ml of acetonitrile and 472 µl of formic acid and adjust volume to 50 ml with deionized water) at 50 psi for 2 min. Then inject the samples at 2 psi pressure for 99 s. Separate small proteins and peptides using a voltage of +25 kV for 30 min at a constant temperature (around 35 °C is initially recommended). Then, in addition to the voltage, separate at 0.5 psi for another 30 min.
5. After each sample run, rinse the capillary with deionized water (50 psi) for 1 min and then with ammonium hydroxide solution (50 psi) for 3 min. Finally, rinse the capillary with running buffer (50 psi) for 3 min.
6. Apply sheath liquid (add 15 ml of 2-propanol and 200 µl of formic acid and adjust the volume to 50 ml with deionized water) coaxially. The running speed of the sheath liquid is defined by the temperature and humidity in the laboratory (normally set at around 20 µl/h). Note that in this protocol no nebulizer gas is applied while the sample is being analyzed.
7. The ESI sprayer must be grounded to achieve electric potential zero and to set the electrospray interface potential at −4000–5000 V.
8. Record MS data at 400–3000 m/z for approximately 60 min, accumulating spectra every 3 s.

LC-MS

The most widely used MS technique today is LC–MS or HPLC–MS. This hyphenated method couples LC (usually a nanoflow HPLC system incorporating a switching valve) and MS, enabling separated biomolecules to be identified by their mass. A standard automated LC-MS system is made up of two three-way diverters, one inline with an autosampler and one switching valve inline with a column, to divert unwanted parts of the eluate to waste before it enters the MS system, the LC system, and the MS. This setup combines the separating power of the LC with the unparalleled capability of the MS to selectively detect and measure molecular parameters such as molecular weight and fragmentation pattern. LC-MS is invaluable for identity confirmation of known analyte molecules, and with modern MS machines it makes possible reconstruction of unknown entities from MS data through qualitative analysis [59]. The following method is a modified version of a published protocol [52].

1. Set up the LC system to consist of sample loop, trap column, and reverse-phase column suitable for nano-HPLC and proteomics analysis.
2. Load 20 µl of each pooled digested protein fraction into a 20 µl sample loop.
3. Inject the sample into the LC system on the trap column and desalt with 3% buffer B (0.1% formic acid in acetonitrile) for 10 min at 5 µl/min.
4. Run a gradient to 80% buffer B and acquire spectra in data-dependent mode in the 400–1800 m/z range. Select the five most abundant +2 or +3 ions in each MS spectrum for MS/MS analysis.
5. Set mass spectrometer parameters at capillary voltage 2.0 kV, capillary temperature 200 °C, and ion trap target value 10,000.

MudPIT

The standard approach for separating complex protein mixtures is 2D-PAGE or its variations. However, its considerable limitations have hampered general implementation of 2D-PAGE in proteomic workflows. These drawbacks include poor separation of highly acidic or basic proteins, a limited molecular mass range, exclusion of the majority of membrane proteins from analysis, gel-to-gel variations, and reproducibility issues. Since many membrane proteins are performing vital functions in cells, it is important to include this class of biomolecules in a global proteomic analysis of cellular systems. Multidimensional protein identification technology (MudPIT) was developed as a gel-free alternative to 2D-PAGE for analysis of both membrane and soluble proteins [60]; it uses column chromatographic separation of biomolecules prior to LC-MS analysis. The following protocol is an adaptation from the literature [61].

1. Chromatography setup is controlled by the Xcalibur interface on the tandem mass spectrometer.
2. Construct a three-phase inline column by packing an end-fitted capillary with reverse-phase resin, strong cation exchange resin, and another layer of reverse-phase resin in sequential order.
3. Load the sample offline onto the column and wash with 500 µl of buffer A (5% acetonitrile, 0.1% formic acid).
4. Install the column assembly in the LC-MS system.
5. Wash the column for 10 min with buffer A and apply a short pulse of water to it (50 µl).
6. Apply a shallow gradient to 84% of buffer B (95% acetonitrile, 0.1% formic acid).
7. Re-equilibrate the column with Buffer A (100%) for 10–15 min.
8. Add 50 µl of 100 mM ammonium acetate using an autosampler. Follow with a shallow gradient of buffer B to 80% acetonitrile, and re-equilibrate with 100% buffer A for 10 min.
9. Repeat step 8 nine times with increasing amounts of 50 µl ammonium acetate (200 mM, 300 mM, 400 mM, 500 mM, 600 mM, 700 mM, 800 mM, 900 mM, and 1 M, respectively).

10. Inject 100 μL 5 M ammonium acetate and elute with a shallow gradient of buffer B to 80% acetonitrile; re-equilibrate with 100% buffer A for 10 min.
11. Analyze by MS and MS/MS and process spectra as usual.

SELDI-TOF MS

Surface-enhanced laser desorption–ionization (SELDI) is a combination of solid-phase chromatography and time-of-flight mass spectrometry (TOF-MS) [62]. Different coatings on the chips allow the binding of various protein classes. ProteinChip arrays are available with different chromatographic properties, including reverse phase (H50), hydrophobic, hydrophilic, weak ion exchange (CM10/WCX2), and immobilized-metal affinity surfaces (IMAC3/IMAC30), or they are preactivated for the coupling of ligands. Thus, each analysis involves an in-built fractionation step, which reduces the sample complexity or enriches target molecules depending on the chip surface.

Sophisticated bioinformatics software tools are used for data processing once SELDI-TOF MS spectra have been collected. They are based on genetic algorithms, decision trees, and a unified maximum separability algorithm [63]. Commonly used are ProteinChip DataManager™, ProteinChip Biomarker Wizard™, and Biomarker Patterns™ (all from Bio-Rad). The following generic protocol is used in SELDI-TOF MS [63]. Note that the minimum sample requirement is 1–10 μg of the total protein per spot and 50–200 μl sample volume if a bioprocessor is used.

1. Assemble arrays in a bioprocessor.
2. Precondition the arrays using an appropriate buffer.
3. Add 50–150 μl of sample to each spot and incubate for 30 min with gentle shaking.
4. Wash three times with 150 μl binding buffer for 5 min with shaking.
5. Wash twice with 150 μl distilled water, 1 min per wash.
6. Remove the bioprocessor; invert and remove any remaining water.
7. Remove the reservoir from the bioprocessor.
8. Allow arrays to air dry.
9. Add twice 1 μl of 50% saturated SPA or CHCA matrix solution to the spot surface and allow to air dry in between.
10. The probes now are ready to be analyzed. If necessary, store the chips in a dark room at room temperature until processing.
11. Place the ProteinChip arrays in the SELDI-TOF MS (PBS IIc ProteinChip reader or ProteinChip System 4000, BioRad) and irradiate with a pulsed UV nitrogen laser.
12. Set the mass analysis range to be evaluated.
13. Adjust the laser intensity and other parameters.
14. Run instrument checks and calibrate the instrument.
15. Save the calibration curves (apply to spectra prior to analysis).
16. Load cassettes with chips into the protein chip reader.
17. Make test shots to optimize laser intensity and detector settings.
18. Incorporate laser intensity and detector settings into a spot protocol.
19. Assemble spot protocols into a chip protocol.
20. Assemble chip protocols into a cassette protocol.
21. Run the cassette protocol and collect data.
22. Normalize the data to a total ion current with an external value of 1.
23. View the spectra for high laser intensities for detection of high-molecular-weight species; set a cut-off mass range around 10 kD to exclude noise and leave the high-mass range unrestricted.
24. View the spectra for low laser intensities for detection of low-molecular-weight species; set a cut-off for exclusion of matrix noise between 1500 and 2000 Da.
25. Set noise definitions to desired settings.
26. Perform manual or automatic peak picking.
27. Perform statistical analysis.

9.2.3 Ionization Techniques

Matrix-assisted laser desorption-ionization (MALDI) and electrospray ionization (ESI) are capable of obtaining complete MS spectra, measuring the presence of molecules in complex mixtures and providing sequence information from biomolecules such as peptides. Both methods are considered “soft” ionization techniques because ions are produced and transferred into a gas phase usually indirectly, thereby enabling measurements such as time-of-flight (TOF). One of the main differences between the two is that ESI produces multiply charged ions, effectively extending the mass range of the analyzer, whereas MALDI generates many fewer multiply charged ions.

9.2.3.1 Electrospray Ionization

Electrospray ionization (ESI) produces a fine mist at the end of a capillary (e.g., chromatography column), which is aimed at an inlet in the mass spectrometer. Liquid is removed via heat and/or collisions with inert gases. Ions are produced by application of a high voltage between the emitter at the end of a capillary and the start of the mass spectrometer [3]. The most practical feature of ESI is that it generates multiply charged ion species [64] and therefore considerably extends the dynamic range. Coupling to a low flow-rate chromatography system can improve the sensitivity of this technique and enables multidimensional online separation and therefore expanded protein and peptide profiling [65]. The following method is an adaptation from the literature [66].

1. Load 10 µL of each pooled digested protein fraction into a 20 µL sample-loop.
2. After injection, trap the sample on a trap column and desalt with 3% mobile-phase B (80% acetonitrile with 0.1% formic acid) for 10 min at 4 µL/min.
3. Start the nano-HPLC gradient elution and acquire spectra in data-dependent mode in the 400–1800 m/z range. Select the five most abundant +2 or +3 ions of each MS spectrum for MS-MS analysis.
4. Alternatively, use a syringe pump to continuously feed the sample.
5. Use common mass spectrometer parameters: 2.0 kV capillary voltage and 150–200 °C capillary temperature.

9.2.3.2 Matrix-Assisted Laser Desorption Ionization

Matrix-assisted laser desorption ionization (MALDI) requires a laser energy-absorbing matrix that rapidly sublimates into the gas phase and transfers some of the energy to an analyte, which also undergoes sublimation. During this process $[M+H]^+$ ions of analyte in the gas phase are formed by ion–molecule reactions [3]. This ionization process necessitates several hundred laser pulses to obtain an adequate signal-to-noise ratio for ion detection; therefore, low shot-to-shot reproducibility and sensitivity to sample preparation methods are the main limitations [67].

MALDI has given rise to a number of derivatives, such as SALDI [68] and DIOS [69], which use porous graphite and silicon instead of a matrix and have been shown to have a higher tolerance toward detergents and salts. AP-MALDI, or atmospheric pressure MALDI, is a hybrid system that allows easy interchange between MALDI and ESI sources [70]. MALDI has also led to the development of SELDI technologies [71], in which a chromatographic functionality is embedded in the plate surface itself. It has also led to the development of material-enhanced laser desorption/ionization (MELDI), in which various derivatized carrier materials are used to batch absorb biomolecules similarly to the SELDI IMAC protocol [72].

The following generic MALDI protocol has been adapted from the literature [48].

1. Redissolve the dried samples in 10 µl of matrix solution—usually either α-cyano-4-hydroxycinnamic acid (CHCA) or sinapinic acid (SPA) dissolved in trifluoroacetic acid acidified acetonitrile.
2. Deposit 0.6 µl of the solution on the MALDI target plate.
3. Allow the sample-matrix mix spots to dry at room temperature.
4. Analyze the samples on a MALDI-TOF/TOF mass spectrometer.
5. Acquire MALDI-MS spectra by averaging 400 individual spectra in positive ion reflector mode at 50 Hz laser frequency at 800–3500 m/z. Perform internal calibration of the MALDI-MS mass spectra using the two trypsin autolysis ions at m/z 842.510 and m/z 2211.105.
6. Obtain MALDI-MS/MS spectra by averaging 800 individual spectra from precursor ions with intensity >1500 in the MALDI-MS spectra.
7. Perform automated analysis of the mass data using, for example, flex Analysis 2.2 software (Bruker) to obtain peptide (MALDI-MS) and precursor (MALDI-MS/MS) m/z values.

9.2.4 Identification of Peptides or Sequencing MS/MS

MS scans usually provide a rich source of information for peak patterns, distribution, and quantitation. However, the identification of peaks by sequence determination requires additional steps, such as collision-induced dissociation (CID) using atom bombardment or energy transfer dissociation (ETD). The fragmentation pattern of such MS/MS scans can be used to mathematically reconstruct a probable amino acid sequence by database matching associated with a protein [73]. A single experiment on a modern mass spectrometer can potentially generate tens to hundreds of thousands of MS/MS scans for thousands of peptides, providing unparalleled efficiency in potential peptide identification [74]. It is therefore required that the sequences be precisely determined with a minimal error rate [75]. Because of the sheer amount of data generated in a typical MS experiment, this process has an absolute requirement of full automation by in silico analysis.

Many software solutions are available to perform MS and MS/MS data processing using sophisticated algorithms and scoring schemes [76]. However, raw data formats are very often vendor-specific and the source data must be converted to a format that can be read by third-party applications while simultaneously retaining as much of the spectral information as possible. This requirement has led to the development of source data standards such as mzXML [77], mzData, and the newer mzML standard [78]. Both PEAKS [79] and MASCOT [80] can accept most MS data formats having to convert the data.

The following is a generic protocol that requires adjustment for a specific instrument [81].

1. Prepare peptide samples by trypsin digest and lyophilize.
2. Prior to analysis, calibrate the instrument using a standard calibration mixture.
3. Check the performance of the instrument with multiple LC-MS/MS analyses of 10 fmol BSA.
4. Resuspend peptides with 30 µL of 0.5% formic acid and incubate for 10 min at 25 °C with shaking.
5. To neutralize the pH, add 15 µL of 0.4 M ammonium bicarbonate and incubate for 10 min at 25 °C with shaking.

6. The resuspended peptides are now ready for LC-MS/MS analysis.
7. Use as mobile-phase A (water with 0.1% formic acid) and mobile-phase B (80% acetonitrile with 0.1% formic acid) for reverse-phase chromatography.
8. Before sample injection, equilibrate the trap column with 100% mobile-phase A and the analytical column with 5% mobile-phase B.
9. Load 40 µL of each peptide sample using an autosampler and desalt on a trap column (0.3 mm id × 5 mm) with a flow of 25 µL/min mobile-phase A for 5 min.
10. Switch the valve to bring the trap column online with the analytical column.
11. Start MS acquisition (for 35 min).
12. Start a gradient elution of the peptides at a flow of 0.3 µL per min on the analytical column, 5–40% mobile-phase B over 30 min, then wash up to 85% B for 5 min, repeat for another 5 min at 85% B, and re-equilibrate at 5% B for 15 min.
13. Use a database search engine such as MASCOT (Matrix Science), OMSSA, or ProteomeDiscoverer (ThermoFisher) with the appropriate reference database.
14. Process MS data using, for example, BioWorks 3.2 (ThermoFisher) to create a peak list file and submit to a MASCOT server for iterative searching on a protein sequence database.
15. Set acetyl (protein N-terminus), carbamidomethyl (C), and oxidation (M) as variable modifications, a peptide mass tolerance of ±20 ppm, and a fragment mass tolerance of ±0.5 Da; allow for a maximum of two missed cleavages.
16. Determine false discovery rates by reverse database searches and empirical analyses of the distributions of mass deviation. MASCOT ion scores can be used to establish score and mass accuracy filters.

9.2.5 Quantitative Approaches

Quantitative proteomics is a crucial technology for understanding biological systems, from protein expression changes to protein modifications associated with molecular pathways, biological processes, and disease conditions. Most quantitative proteomic experiments use isotopic or isobaric labeling strategies, which can then be differentiated by MS analysis. Relative quantitation methods (SILAC, ICAT, and iTRAQ) are generally used for comparisons of protein or peptide abundance levels between samples, while spiking unlabeled samples with known concentrations of isotopically labeled synthetic peptides can result in absolute quantitation of target peptides (SRM, MRM). Label-free methodologies are also useful for both relative and absolute quantitation, though they are profoundly more complex than protein and peptide identification.

9.2.5.1 ICAT

The isotope-coded affinity tag (ICAT) method was originally developed to reduce sample complexity and identify low-abundance proteins and peptides in complex samples [8]. ICAT tags originally consisted of a sulfhydryl-reactive chemical cross-linker, which can react with free thiols on cysteine residues, an 8-fold deuterated (d8) or light (d0) linker region and a biotin group. The proteins are then passed over immobilized avidin, which binds to the biotin tag and thereby purifies the labeled proteins from the sample. After this purification step, heavy (d8) and light (d0) samples are combined and analyzed for relative quantitation by LC-MS and identification by MS/MS.

Because cysteine residues are not present in all proteins, this methodology does not result in global labeling and therefore is an intrinsic approach to reduce sample complexity. However, this can also pose a problem if the target proteins contain very low levels of available reactive groups. Therefore, the difference in amino acid availability should be considered when determining the right isotopic labeling method quantitative proteomic analyses [82]. Modified ICAT tags were also developed that contain ^{13}C instead of deuterium to avoid the problem of partial peak partitioning during LC separation [83]. The following protocol is an adaptation of a previously published application [84] that uses the ICAT labeling kit (Invitrogen).

1. Denature the sample by incubation at 95 °C.
2. Add a reducing agent such as DTT to a final concentration of 5 mM and incubate at 37 °C for 20 min. Precipitate the proteins by adding 6 to 7 volumes of cold acetone and incubate at −20 °C for at least 1 hour. Spin the sample, dry the pellet, and resuspend in a minimal amount of 5 mM Tris–HCl, pH 8.3, 10 mM EDTA.
3. Dissolve the ICAT reagent in 20 µl of acetonitrile per tube. Add the sample to the ICAT tube at a minimum concentration of 6-fold molar excess of ICAT reagent. It is important that the ICAT reaction be conducted in as small a volume as possible, ideally less than 150 µl per tube.
4. Incubate the reaction at room temperature in the dark for a minimum of 2 h.
5. Check the ICAT label incorporation on SDS–PAGE, which will result in a shift in the protein bands.
6. Stop the reaction by adding 5 mM DTT and incubate at 55 °C for 20 min.
7. Combine the ICAT-labeled samples at the desired ratio, ensuring that the final urea concentration is less than 1 M.
8. Trypsin-digest proteins at a 1:50 (w/w) ratio (trypsin/sample) overnight at 37 °C.
9. Remove excess unreacted ICAT tags by cation exchange separation prior to avidin purification.

10. Clean the ICAT-labeled peptides by avidin column chromatography according to the manufacturer's protocol.
11. Cleave off the biotin moiety by adding 95 μl cleaving reagent A and 5 μl cleaving reagent B to each avidin-purified fraction and incubate at 37 °C for 2 h.
12. Quench the reaction by adding 8 μl 0.2–0.4% acetic acid to each fraction, mix and centrifuge. Transfer the supernatant to a new vial for LC-MS/MS analysis.

9.2.5.2 iTRAQ

iTRAQ uses isobaric tags, originally termed tandem mass tags, that have identical masses and chemical properties; therefore, heavy and light isotopologs co-elute during LC separation. These tags are subsequently removed from the peptides during the LC-MS/MS run by CID [10]. A clear advantage of these isobaric mass tags is the flexibility they provide in choice of labels as well as their multiplex capabilities. Commercially available tags—for example, TMT™ or iTRAQ™—offer the simultaneous analysis of 4, 6, or 8 biological samples. The basic components of all isobaric mass tag reagents are (1) a mass reporter (tag) containing a unique number of ^{13}C substitutions and (2) a mass normalizer with a unique mass used to make all tags equal in mass. Additionally, isobaric mass tags possess a reactive moiety that cross-links to a specific chemical group in a polypeptide—for example, primary amines or cysteines—and are designed to fall apart at a specific linker region during high-energy collision-induced dissociation (HCD). These different-sized tags are then analyzed by LC-MS/MS, resulting in concurrent peptide identification and relative quantitation. Furthermore, because MS/MS is required to detect the isobaric tags, unlabeled peptides cannot be quantitated. The following protocol is an adaptation of a previously published method [85] and uses the iTRAQ labeling kit (Invitrogen).

1. Extract proteins from tissues, cultured cells, or biofluids in a homogenization buffer of choice (200 μl) and sonicate for 2 min.
2. Spin at 12,000 × g to pellet insoluble materials.
3. Determine the total protein concentration.
4. (Optional) Clean up the sample by acetone precipitation: add 6 volumes of cold acetone to the cold sample tube, mix and incubate at −20 °C until a precipitate forms (about 1 h); then spin the sample to pellet the proteins, decant the acetone, and air dry.
5. To each tube containing 5–100 μg of sample, add 20 μl of a resuspension buffer of choice.
6. Add 1 μl of the denaturant solution with a final SDS concentration of 0.1%, and mix by vortexing.
7. To each sample, add 2 μl reducing reagent, mix, and incubate at 60 °C for 1 h; then spin and add 1 μl cysteine-blocking reagent, mix, and incubate at room temperature for 30 min in the dark.
8. Add 10 μl of the trypsin solution (final concentration of 5–10 ng/ml), mix, and incubate overnight at 37 °C.
9. Add the iTRAQ™ reagents, which are dissolved separately in 70 μl ethanol, to different samples (e.g., iTRAQ Reagent 114 to sample 1 and iTRAQ Reagent 117 to sample 2), vortex, and incubate at room temperature for 1 h.
10. Combine the contents of each iTRAQ reagent-labeled sample tube into one tube and mix, following with incubation at room temperature for 2 h; quench the reaction with 3 μl glacial acetic acid once the labeling is complete. Combine all labeled samples.
11. Clean up the sample to remove excess iTRAQ reagents, salts, and other interfering compounds. This can be done by batch absorption onto cation exchange chromatography resins, reverse-phase resins, or acetone precipitation.
12. Dry the sample on a Speedvac.
13. Dissolve samples in LC-MS buffer A and transfer to injection vials for MS/MS analysis.
14. Analyze MS data using, for example, Proteome Discoverer with MMTS-labeled cysteines and iTRAQ-labeled N-termini as fixed modifications, and phosphorylation of serine, threonine, and tyrosine and iTRAQ labeling of threonine and lysine residues as variable modifications for the search parameter settings. Alternatively, many iTRAQ analysis tools are freely available, such as LTQ-iQuant [86], IsobariQ [87], OCAP [88], MilQuant [89], PeakQuant [90].

9.2.5.3 SILAC

SILAC is based on the comparison of tagged and untagged peptides by MS using a metabolic labeling strategy. This is achieved by growing cells in an environment where essential components, such as amino acids, are supplemented in either heavy or light isotopic forms. There are several methods of this type of *in vivo* labeling, and selection criteria include the extent of labeling required. One traditional way is to grow cells in an environment where the only nitrogen source is ^{15}N-labeled ammonium persulfate [91], resulting in uniformly labeled amino acids containing heavy nitrogen (^{15}N) only.

A further development and adaptation for mammalian cell cultures is the method for stable isotope labeling by amino acids in cell culture (SILAC), which is nowadays the most common approach for *in vivo* isotopic labeling [92]. This is performed by growing cells in a culture medium that contains $^{13}C_6$-lysine and/or $^{13}C_6$-arginine. These two amino acids are selected since trypsin, the main enzyme used in proteomics to generate proteolytic peptides for MS analysis, cleaves at the C-terminus of lysine and arginine residues. Therefore, all tryptic peptides from cultures grown in SILAC media, with the exception of the extreme C-terminal peptides, contain at least one labeled amino acid.

The key aspect is that, since heavy and light samples are combined prior to sample preparation for MS analysis, the influence of processing errors is minimized, resulting in a much more robust quantitation analysis. This makes this method particularly useful in the detection of small changes in protein levels or PTMs between two to three sets of conditions. Another benefit is that metabolic labeling strategies can yield more than 90% isotopic incorporation in immortalized cell lines [93], again resulting in an improved quantitation result.

The downside of metabolic labeling is that it is not generally compatible with any cell line, as it depends on a cell culture cycle that lasts several days, and the conditions used to incorporate the isotopic tags might also influence the cells directly [10]. Additionally, because of the limited number of heavy isotopes incorporated into lysine and arginine—namely, unlabeled, $^{13}C_6$-, and $^{13}C_6$ $^{15}N_4$-labeled amino acids—in any SILAC experiment, a maximum of three conditions can be studied simultaneously. The following is an adaptation of a previously published protocol [84].

1. Prepare SILAC media. For example, add normal methionine to both "light" and "heavy" medium if SILAC labeling with arginine and lysine only, and add appropriate amounts of "light" and "heavy" arginine and lysine to the "light" and "heavy" media, respectively.
2. Grow cells in "light" or "heavy" SILAC medium according to standard cell culture conditions.
3. Lyse cells by adding 300 µl of a solution containing 6 M urea and 2 M thiourea to a cell pellet or a cell culture dish; then vortex and pellet the debris by centrifuging for 10 min at 16,000–20,000 × g at 18 °C.
4. Collect the supernatant in a new tube, avoiding DNA and the cell pellet.
5. Measure the protein concentration using a standard protein assay and use about 25–50 µg of protein for the in-solution digest.
6. Reduce disulfide bonds by adding DTT to a final concentration of 1 mM and incubate at 37 °C for 30 min.
7. Alkylate cysteins by adding iodoacetamide to a final concentration of 5 mM, mix, and incubate in the dark for 20 min.
8. Quench the reaction by adding DTT to a final concentration of 5 mM.
9. Reduce the final concentration of urea to 3 M by adding an equal volume of 50 mM ammonium bicarbonate.
10. Add endoproteinase Lys-C at an enzyme-to-substrate ratio of 1:100 and incubate at 37 °C for 2 h.
11. Reduce the urea concentration to less than 1.5 M by diluting the lysate with an equal volume of 50 mM ammonium bicarbonate.
12. Add trypsin at an enzyme-to-substrate ratio of 1:50 and incubate at 37 °C for 4 h or overnight.
13. Acidify with trifluoroacetic acid, 1% final concentration, to stop digestion.
14. Analyze the sample by nanoflow LC–MS/MS to identify proteins and peptides.
15. Determine the degree of incorporation by looking for "heavy" and "light" peptides using MS.
16. Analyze MS data using specific SILAC software such as PeakQuant [90], IsoQuant [94], or SILACtor [95].

9.2.5.4 SRM, MRM, and label-free

Selected reaction monitoring (SRM) and multiple reaction monitoring (MRM) are methods of absolute quantitation, also referred to as AQUA. Both are targeted proteomics analyses carried out by spiking complex samples with stable isotope-labeled synthetic peptides that act as internal standards for specific peptides of proteins to be investigated [96]. These labeled peptides are designed to be identical to peptides observed in the samples to be investigated—for example, tryptic peptides generated by sample digestion—so that they co-elute with the target peptide and are simultaneously analyzed by MS/MS. Predetermined calibration curves of the heavy peptides are then used to calculate the concentration of the target peptides to a high accuracy. However, this means that for each target peptide in the sample, a calibration curve has to be generated. The drawback of these techniques is that they are considerably labor-intensive and can also be economically unfavorable.

Rapid and low-cost alternatives compared to other quantitative proteomic approaches have been developed over the last years. These label-free methods are usable for both relative and absolute quantitation and are ideal for large-scale sample analyses in clinical screening or biomarker discovery experiments [97]. The main issue is that they are less reliable to measure small changes and can have a limited range of linear quantitation [98]. However, improved MS instruments and workflows are slowly filling this gap. Protein quantitation is done using either spectral counting or ion peak intensity [99].

Only LC-MS scans are used for relative quantitation by ion peak intensity. The signal intensities of all m/z peaks across all scans are used to determine relative quantities, since these values have been shown to correlate highly with ion concentrations [100]. Because of the considerable amount of data collected from such experimental runs, sensitive computer algorithms are needed for automated ion peak alignment and comparison.

Label-free relative quantitation by spectral counting compares the sum of MS/MS spectra of a peptide across multiple samples; it has been shown to correlate directly with protein abundance [100]. Unlike peak intensity quantitation, special algorithms or other tools are not required for spectral counting, but significant normalization is necessary [101].

Apart from relative quantitation, label-free methods can also be used to determine the absolute concentration of proteins in a sample. The exponentially modified protein abundance index (emPAI), which estimates protein abundance based on the number of detected peptides and the number of theoretically observed tryptic peptides for each protein, determines the approximate absolute protein abundance in large-scale MS experiments [97]. Another method, absolute protein expression (APEX), is based on spectral counts and uses correction factors to adjust protein abundance proportional to the number of peptides observed [102].

Generation of an SRM method

SRM studies are by default hypothesis-driven and focus on a specific set of proteins that correspond to a specific clinical or biological question, such as a biomarker, a biochemical pathway, or a protein network [103].

The critical part of the experimental design is the selection of peptides that will act as surrogates for the proteins of interest. In order to accomplish this, two to three peptides are first selected to perform the quantitation for each protein of interest. These peptides should conform to the expected peptides to be analyzed—for example, proteotypic if the samples were proteolytically processed—and they should be unique to the protein of interest. For previously analyzed proteins in proteomic discovery experiments, the unique peptides observed represent the starting point of the targeted SRM method design. They can be directly imported in specific software packages for SRM method design [104]. Existing MS/MS spectral libraries are queried to extract fragment ions to generate the SRM transitions, usually the most intense assigned fragment ions. If the proteins have not been reported before, the peptide selection is performed in silico using the protein sequences.

The peptides should ionize well in ESI mode and produce doubly or triply charged ions in the mass range of 400–1500 m/z, corresponding in length to peptides of 8–20 amino acid residues. For reasons of practicality, peptides containing residues such as methionine, which can undergo easy modification, are excluded.

Precursor and fragment ion m/z are calculated for each peptide, and the reverse-phase elution time is predicted using hydrophobicity indexes [105]. For each peptide, a series of transitions are defined. If the MS/MS spectra of the peptides are available, five to eight of the most intense y ions are chosen from the reference spectrum. If no MS/MS information is available, the m/z values of the fragments are calculated, and y fragments with m/z values above the precursor ion are selected.

The collision energy in a preliminary study is calculated for each peptide based on the generic formula of collision energy in volts $= 2 + 0.034 \times$ precursor ion m/z. This transition list is exported and uploaded into the instrument method.

The samples are spiked with various concentrations of the selected peptides, and the mixes are analyzed by targeted LC-MS/MS. The SRM transitions are monitored for endogenous and standard peptides.

Design of MRM Transitions/MRM Assays

Generally, MRM analyses can be carried out on any triple quadrupole MS instrument. However, the specific details of parameter optimization and MRM assay setup can differ between instruments [106]. Numerous software solutions exist for both design and analysis of MRM assays [107]. The following protocol can be used.

1. Obtain standard peptides relating to the proteins to be analyzed and dilute each one with 40% (v/v) acetonitrile, 0.1% (v/v) FA, to a final concentration of 200 fmol/μl.
2. Analyze each standard peptide by nanoSpray direct-infusion MS (ESI) to generate MRM transitions.
3. Measure the peptides by enhanced-resolution single MS and enhanced product ion MS/MS scans to obtain molecular weight and fragmentation patterns.
4. Select the monoisotopic m/z of the doubly or triply charged precursor as the reference mass, and choose the three or four most abundant fragments with an m/z higher than that of the precursor.
5. Optimize the parameters that show the highest signal intensity by adjusting the declustering potential, the entrance potential, the collision energy, and the collision cell exit potential.
6. Using the expected mass differential from the standard peptides, calculate the MRM transitions for the corresponding endogenous peptides from the samples to be analyzed.
7. Check for interfering signals by analyzing the standard peptides and the endogenous peptides separately.
8. Dilute the standard peptides to a concentration of 2.5 pmol/μl with 2% (v/v) acetonitrile, 0.5% (v/v) formic acid, and prepare a peptide mixture containing all standard peptides at 100 fmol/μl each in the same buffer.
9. Mix the proteolyzed sample with standard peptides at the desired concentrations; use combinations of different concentrations for both standard peptides and samples.
10. Analyze the sample containing endogenous and standard peptides by targeted LC-MS/MS; monitor the MRM transitions for endogenous and standard peptides.

Label-Free Using Ion Intensities

The number of ions in a peak from a mass spectrum, at a particular m/z ratio, is reflected by the height or area of that m/z detected by the mass spectrometer at any given time.

This is generally known as the ion abundance. The ion abundance cannot be used directly to infer absolute protein or peptide quantities, since each peptide has different ionization efficiencies. However, by comparing the ratio of ion abundances between identical peptides obtained in different experimental runs, it is possible to estimate a differential expression [99]. Here is the protocol.

1. Search all acquired spectra using a database search algorithm to identify all proteins/peptides.
2. Extract the ion abundances for all charge states for each identified peptide and, where necessary, combine them for each unique peptide from each sample and calculate an abundance ratio across all samples.
3. Calculate protein abundance ratios by combining all peptide ratios.
4. Normalize the protein abundance ratios by calculating the mean of all protein ratios, and divide all ratios by this value.
5. Determine the statistical significance of all protein abundance ratios across all samples, including the correlation and covariance values.

Label-Free Using Spectral Counting

A protein's spectral count refers to the total number of MS/MS spectra acquired from proteolytic peptide ions for that specific protein during a LC-MS/MS run. This analysis method is based on the abundance of a peptide, so the more abundant the peptide, the more likely it will be selected for MS/MS analysis. In several comparison studies, it was shown that the correlation of protein abundance with spectral count is better than that which can be obtained from protein sequence coverage or peptide count [100].

1. Search all acquired spectra using a database search algorithm to identify all proteins.
2. For each protein, count all MS/MS spectra for each peptide belonging to the same protein.
3. Normalize the total spectral count across all samples. The total spectral count should be the same under identical data acquisition conditions. Any one sample can be used as the reference value.
4. Determine the statistical significance of protein spectral counts across all samples. A number of different statistical tests such as t-test or Mann-Whitney can be used.

9.3 DATA ANALYSIS

State of the art high-throughput MS-based proteomics methodologies require a considerable amount of data processing, management, and visualization. These are central and critical components (Figure 9.3), and are often some of the most time-consuming steps [108].

The development of better MS machines, technology, and strategies has led to a wider use of proteomics and to a tremendous increase in the development of new software tools, including freely available and open-source programs and scripts, ranging from database search analysis to post-processing of mass spectra and resulting data. Table 9.3 is an overview of selected software libraries, both commercial and open-source, and information on some of the available applications.

9.3.1 Downstream Processing

9.3.1.1 Normalization

Technical and biological variations can lead to deviations in quantified MS peak intensities, resulting in unreliable run-to-run reproducibilities. Therefore, normalization of peak intensities across an MS proteomics dataset is a fundamental step in processing MS data [124]. Furthermore, it is crucial to select the appropriate normalization method, which will have a considerable impact on the downstream analysis. Normalization procedures are generally included and sometimes automatically applied in the software used to process the raw MS data. They must be recalculated for each independent dataset and be independent from the biological factors associated with the experiment.

9.3.1.2 Statistical Analysis

Discovery of biomarkers and establishing biomarker models usually require a statistical analysis of peak patterns, as does feature selection to decide if proteins are promising leads for biomarker candidate selection [125]. The classification rules use information about the contribution to the classification of each variable (e.g., peak intensity) which can then lead to the mathematical identification of biomarkers. Two methods that determine the interesting variables directly are the classification tree [126], which classifies samples based on their values for a small number of proteins, and nearest shrunken centroid (NSC), a progressive sampling algorithm [127]. Linear support vector machines (SVM) use relevant information about variables and features in the form of weights and regression coefficients. This type of analysis is used in many applications to select relevant sets of biomolecules.

Rank products were initially designed for gene selection using gene expression differences between two groups directly [128], but they have also been employed for selection of proteins using a principal component discriminant analysis classification rule [129]. Another classification method is PLSDA, a multivariate inverse least squares discrimination method used to classify features; thresholding of regression coefficients using PLSDA was used to identify potential metabolite biomarkers [130]. A variant of PLSDA is variable importance in the projection (VIP).

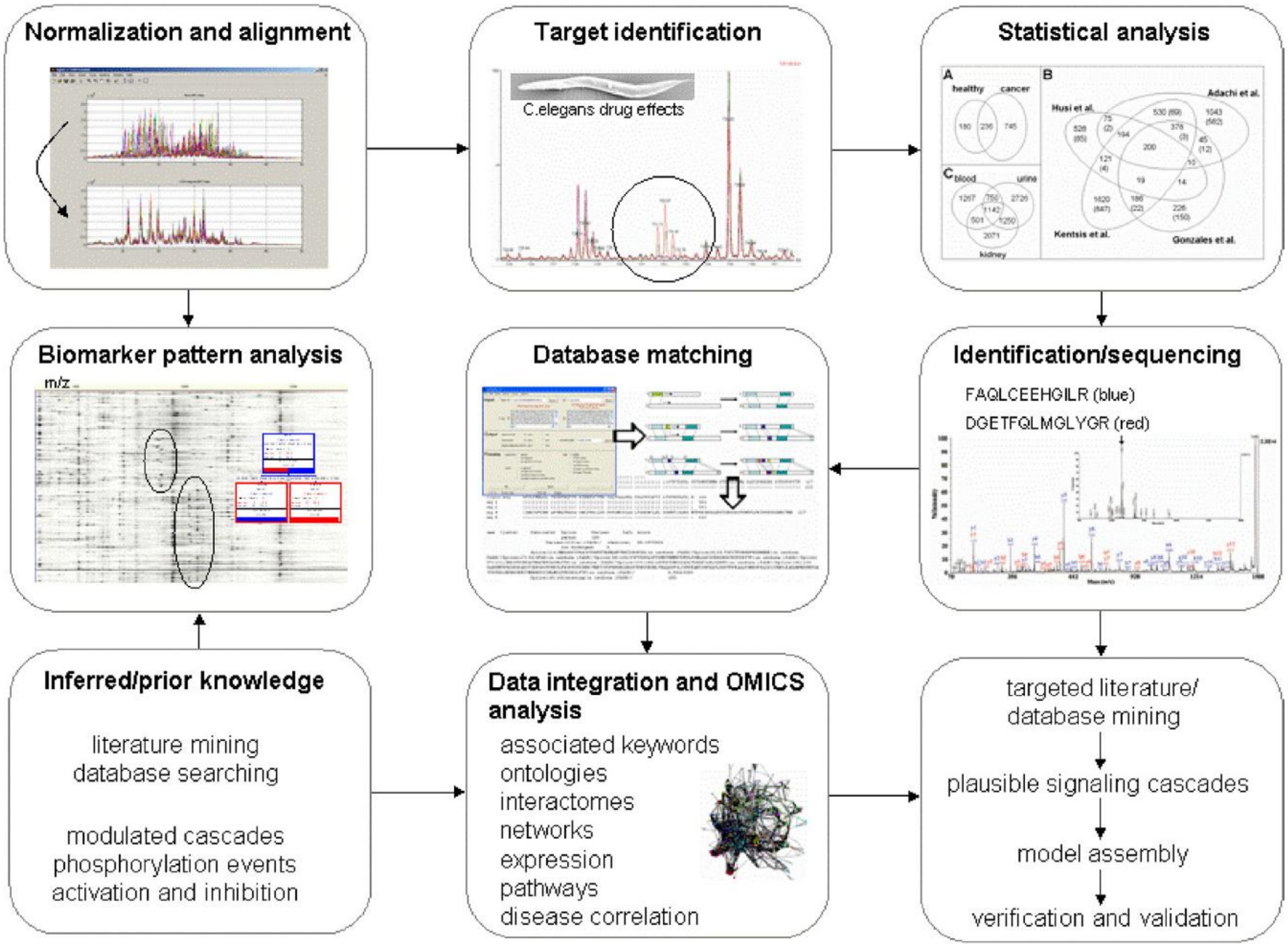

Normalization and alignment
Target identification
C.elegans drug effects
Statistical analysis
A
healthy
cancer
B
Husi et al.
Adachi et al.
C
blood
urine
kidney
Kentsis et al.
Gonzales et al.
Biomarker pattern analysis
m/z
Database matching
Identification/sequencing
FAQLCEEHGILR (blue)
DGETFQLMGLYGR (red)
Inferred/prior knowledge
literature mining
database searching
modulated cascades
phosphorylation events
activation and inhibition
Data integration and OMICS analysis
associated keywords
ontologies
interactomes
networks
expression
pathways
disease correlation
targeted literature/
database mining
plausible signaling cascades
model assembly
verification and validation

FIGURE 9.3 **Data analysis workflow.** Proteomics data processing often requires several steps and more than one software tool; the most common procedures are indicated according to the aim of the analysis.

The VIP value of a variable reflects its importance in the model with respect to the response vector as well as to the projected data [131].

Variable selection using multiple classification rules is computationally more demanding because of the amount of information that has to be permuted. The random forest algorithm, or decision tree analysis, which is typically used in SELDI-TOF-MS biomarker model building, estimates the importance of a given variable by mathematically rearranging its measurements, leaving the rest of the data intact and classifying new samples [132]. Significance tests, such as t-tests and Wilcoxon-Mann-Whitney tests, can also be used to select disease markers, without running a classification algorithm [133].

9.3.1.3 *Identification*

Nowadays, there are a considerable number of diverse mass spectrometers used to generate MS data in proteomics. This is also the case for algorithms employed in querying and cross-comparing tandem MS data [134]. The most popular software packages used in protein identification from raw MS scans are Sequest, MASCOT, OMSSA, and X!Tandem [135]. However, further developments and the emergence of new tools such as Morpheus [136], as well as specialized tools like MaxQuant [137], which is often used for iTRAQ quantitation, will accelerate identification and enhance feature extraction.

9.3.2 Biomarker Pattern Analysis

A diagnostic-proteomics approach does not necessarily rely on the identification of peaks to be useful in diagnosing a disease condition; on the other hand, a biomarker-discovery approach usually does. If, however, these peaks of interest, which show a difference in abundance according to disease states, are subsequently identified and, for example, validated by independent techniques such as Western blotting, it is possible to learn more about biological context and potential impact on a biological system. This diagnostic-proteomics approach has become popularly known as proteomic-pattern analysis [138]. Pattern matching was used

TABLE 9.3 Analysis Tools

Software	Area	Usage	Source	Availability	Reference
Univariate statistics	Statistics	Statistical tools to calculate covariances, confidence values, etc.; well established and easily interpretable; does not work with multiple variables	—	Many software packages available	[109]
PCA	Statistics	Principal component analysis to calculate and visualize potential correlations of datasets; good visual presentation	—	Many software packages available	[110]
Random forest/ decision tree/ support vector machine	Statistics	Classification and regression analysis to build decision models (biomarker models/patterns); random forest controls overfitting through subset selections; number of incorrect predictions increases with decision tree complexity; support vector machine shows poor classification of blinded data with high numbers of variables and dimensions because of overfitting	—	Many software packages available	[111]
Sequest/ X!Tandem/ MASCOT	MS analysis	Mass spectrometry search engine to identify and extrapolate peptide/ protein sequences	http://proteomicsresource.washington.edu/sequest.php; www.thegpm.org/tandem/; www.matrixscience.com/	Commercial and open-source	[112]
KEGG	Pathway analysis	Collated metabolic and signaling maps from the literature	www.genome.jp/kegg/	Commercial	—
Ingenuity pathway analysis (IPA)	Pathway analysis	Large-scale pathway analysis; simple user interface, easy to use	www.ingenuity.com/	Commercial	—
Reactome	Pathway mapping	Represents many diverse processes in a biological system, including pathways of intermediary metabolism, regulatory pathways, and signal transduction	www.reactome.org	Open-source	[113]
PathVisio	Pathway analysis	Pathway display and analysis based on vector graphics, "classical" pathway diagrams, manually expandable maps, statistical analysis	www.pathvisio.org	Open-source	[114]

(*Continued*)

TABLE 9.3 Continued

Software	Area	Usage	Source	Availability	Reference
CytoScape	Multiple applications	Network data integration, analysis, and visualization; generic visualization tool; many third-party plug-ins available; nodal display of large networks can be difficult	www.cytoscape.org/	Open-source	[115]
Blast	Sequence analysis	Global sequence searching; most powerful, general tool of choice; computationally intensive	http://blast.ncbi.nlm.nih.gov/Blast.cgi	Open-source	[116]
Clustal	Sequence analysis	Discrete sequence alignments; unsuitable for multiprotein cluster analysis	www.clustal.org/download/current/	Open-source	[117]
LSCluster	Sequence analysis	Sequence clustering, aligning, whole proteome analysis; deprecated sequence alignments; lacks heuristics and substitution tables	www.padb.org/download/download.html	Open-source	[118]
ClueGO	Gene ontology analysis	Visualizes nonredundant biological terms for large clusters of genes in a functionally grouped network	http://apps.cytoscape.org/apps/cluego	CytoScape plug-in	[119]
MiMI plug-in	Interactome mapping	Analyzes molecular interactions from the Michigan Molecular Interactions (MiMI) database	http://mimi.ncibi.org	CytoScape plug-in	[120]
GeneMania	Network analysis	Identifies the most genes related to a query gene set using a guilt-by-association approach	www.genemania.org/plugin/	CytoScape plug-in	[121]
NetworkAnalyzer	Network analysis	Analyzes biological networks and calculates network topology parameters, including diameter of a network, average number of neighbors, and number of connected pairs of nodes	http://apps.cytoscape.org/apps/networkanalyzer	CytoScape plug-in	[122]
OmicsAnalyzer	Network analysis	Integrated analysis for large-scale omics data	http://vrac.iastate.edu/~jlv/omicsanalyzer/	CytoScape plug-in	[123]

successfully in comparing a biomarker profiling study with lists of proteins identified in the same samples, followed by cross-matching of observed proteins and expected peaks using a specific pattern-matching algorithm [139]. Other pattern-matching approaches are used in 2D-gel analysis [140] and LC-MS screens using MassUntangler software [141].

9.3.3 Data Integration and OMICS Analysis

Systems biology, or integrative biology, uses descriptive terms to understand complex interactions based on the analysis of network behavior and dynamic aspects. It combines proteomics, metabolomics, and transcriptomics/genomics data with a systems approach to investigate a clinically relevant complex pathophysiology [142]. Thus, it identifies key signaling cascades [143], inferred metabolic alterations [144], and molecules modulated in highly complex diseases [145]. Apart from the pure gain of knowledge, such holistic multi-OMICS approaches have become increasingly important in healthcare, as they are considered key technologies for personalized and preventive medicine [146]. The addition of chronological and spatial aspects enables mathematical modeling [147]—for example, in signaling or gene activation cascades. This can lead to novel insights into the mechanisms of pathogenesis and can support the development of new therapeutics [148]. Multi-OMICS approaches can also be used to discover novel biomarkers for disease diagnostics [149].

In practical terms, systems biology is currently one of the more demanding scientific disciplines because of its lack of all-encompassing software tools, disparate data sources, data structures, and formats, demanding data processing and interpretation of results as well as skills to judge experimental findings in light of existing knowledge obtained from the scientific community [150].

9.4 DISCOVERY AND VALIDATION

Proteomics has been enthusiastically taken up by the scientific community in all areas of life sciences, and there is an ever-increasing number of studies using proteomic workflows, taking advantage of improved instrumentation, methodologies, and analysis software [168]. Table 9.4 lists a representative number of investigations using the methods mentioned in this chapter to define new and usable biomarkers that might be of clinical benefit in the future.

Additionally, because the level of biological information is also on the increase, the option of looking not just at a specific set of molecules, but rather at the entire biological system, becomes a workable reality. This approach, termed systems biology analysis, is ideally suited to investigating and describing whole-systems perturbations based on disease conditions; it has been applied to several disease models to date, of which a selection is listed in Table 9.5. The following paragraphs focus on diseases and biological systems using proteomic techniques that have garnered particular interest.

9.4.1 Cancer Markers

Undoubtedly, one of the major areas of biomarker investigation using proteomics techniques is cancer research. Numerous studies have been carried out; however, the usability of results as biomarker-screening tools in the clinical setting is relatively low. To a lesser extent, the use of MS for discovery of proteomic biomarkers as a cancer diagnostic tool, using various analytical platforms, has been reported [139].

One of the main applications of MS-based analyses is screening and detection of cancer in its early stages, as well as prediction of cancer development [176]. Urine as a source of screening medium has gained considerable interest because of its noninvasiveness and biomolecular stability compared to serum. Proteomic urinary biomarker screening of bladder cancer samples has shown several proteins to be differentially present—namely, bikunin and pro-urolinase-plasminogen activator [177], apolipoprotein A-I, apolipoprotein A-II, heparin cofactor 2 precursor, and peroxiredoxin-2 [178], or PLK2 [179]. In contrast, ovarian cancer biomarkers have been described as fragments of fibrinogen alpha, collagen alpha-1, and fibrinogen beta-NT [180].

Metzger et al. [181] used CE-MS to establish a biomarker model based on 42 urinary peptides for the diagnosis of cholangiocarcinoma with a sensitivity of 83% and a specificity of 79%. The majority of peptides identified in this study were fragments of interstitial collagens; immunostaining of liver sections for matrix metallopeptidase 1 showed an increased expression of interstitial collagenase in liver epithelial cells of cholangiocarcinoma patients. This presented a possible explanation for the observed altered collagen fragmentation. SELDI-TOF-MS was also used extensively in the discovery of cancer biomarkers.

A study by Husi et al. [139] used a combination of SELDI-TOF-MS, LC-MS/MS, and pattern matching to identify a urinary biomarker profile of six peptides in upper gastrointestinal cancer with a sensitivity of 96% and a specificity of 72%. The identified molecules consisted of S100-A6, S100-P, S100-A9, S100-A12, ITIH4, and ubiquitin. Several of these were validated by Western blotting.

9.4.2 Cardiovascular Diseases

Stroke and heart failure rates are globally on the increase; these severe diseases are associated with both poor prognosis and adverse outcomes [182]. Several studies have sought to uncover the molecular pathways involved using various heart failure models and proteomic techniques, including

TABLE 9.4 Proteomics in Biomarker Research

Disease Area	Disease	Tissue	Platform	Identified Biomarker	Number of Modulated Proteins/ Models/Peak Patterns	Reference
Fibrosis	Cystic fibrosis	Nasal cells	2DE, MALDI	—	65 biomarkers	[151]
Cancer	Esophageal carcinoma	Laser capture of microdissected tissue	2D DIGE, MALDI	Annexin I, gp96	165 biomarkers	[152]
Cancer	Hepatocellular carcinoma (HCC)	Serum	2DE, LC-MS/MS	Complement C3, APOA1		[153]
Toxicity	Arsenic exposure	Mouse lung-lining fluid	MudPIT	RAGE		[154]
Cancer	Pancreatic cancer	Cancer cells	MudPIT	Versican, perlecan, syndecan 1 and 4		[155]
Cardiovascular	Heart failure (HF)	Serum	2D DIGE, LC-MS/MS	Leucine-rich α2-glycoprotein (LRG)		[156]
Cardiovascular	Atherosclerosis	Secretome	LC-MS/MS	Gelsolin, vinculin, lamin A/C, phosphoglucomutase 5		[157]
Cancer	Ovarian cancer	Serum	MALDI	APP, CA 125, CCL18, CXCL1, IL-8, ITIH4		[158]
Cancer	Gastric adenocarcinoma	Serum	SELDI, LC-MS/MS	Fibrinogen α-chain, apolipoprotein A-II, and apolipoprotein C-I		[159]
Cancer	Upper gastrointestinal cancer (UGI)	Urine	SELDI, LC-MS/MS	S100-A6, S100-P, S100-A9, S100-A12	Peak pattern, 6 biomarkers	[139]
Cardiovascular	Coronary artery disease (CAD)	Urine	CE-MS	SERPINA1, COL1A1 COL3A1, PCSK1N, PGRMC1, FXYD2, FGA	Peak pattern, 238 biomarkers	[160]
Neurological	Alzheimer's disease (AD)	Cerebrospinal fluid (CSF)	CE-MS	ProSAAS, apolipoprotein J, neurosecretory protein VGF, phospholemman, chromogranin A	Peak pattern, 12 biomarkers	[161]
Neurological	Multiple sclerosis	Cerebrospinal fluid (CSF)	iTRAQ, SRM	α-1-antichymotrypsin, contactin-1, apolipoprotein D, clusterin, kallikrein-6	11 differentially expressed proteins	[162]

TABLE 9.4 Continued

Disease Area	Disease	Tissue	Platform	Identified Biomarker	Number of Modulated Proteins/Models/Peak Patterns	Reference
Cancer	Glioblastoma	Plasma	iTRAQ	FTL, S100A9, CNDP1	61 differentially expressed proteins	[163]
Neurological	Parkinson's disease (PD)	Cerebrospinal fluid	iTRAQ, MRM	Tyrosine-kinase-non-receptor-type 13, Netrin-G1, Golgin-160, Apolipoprotein B-100	16 differentially regulated proteins	[164]
Cancer	Melanoma metastasis	Cell culture	SILAC	CDCP1		[165]
Development	Preeclampsia (PE)	Serum	SRM	AHSG, RBP4, AMBP, IGFBP-ALS		[166]
Cancer	Liver cancer	Serum	MRM	Vitronectin, clusterin		[167]

TABLE 9.5 Proteomics and Systems Biology Analysis of Systems Perturbations

Disease Area	Disease Model	Tissue	Platform	Target Identification	Modulated Proteins/Models	Pathways	Reference
Cancer	Prostate cancer	Micro-dissected tissue	2-D DIGE, MALDI	—	79 differentially expressed molecules	Androgen receptor signaling pathway, cell differentiation, apoptosis, protein metabolic processes, Ca^{2+} signaling pathways	[169]
Cancer	Ovarian cancer	Cancer cell lines	MudPIT	—	196 differentially expressed proteins	Actin cytoskeleton, extracellular matrix (ECM) and focal adhesion pathways	[170]
Cancer	Glioblastoma multiforme (GBM)	Cerum	2-D DIGE, MALDI	Haptoglobin, plasminogen precursor, apolipoprotein A-1 and M, transthyretin	55 differentially expressed proteins	Intrinsic prothrombin activation pathway, plasminogen-activating cascade, coagulation system, glioma invasiveness signaling, PI3K signaling	[171]

(*Continued*)

TABLE 9.5 Continued

Disease Area	Disease Model	Tissue	Platform	Target Identification	Modulated Proteins/ Models	Pathways	Reference
Cardiovascular	Chronic kidney disease (CKD), type 2 diabetes mellitus (T2DM)	Serum	iTRAQ, MRM	—	—	Lipid and small molecule metabolism, pathways that lead to disruption of glucose homeostasis and blood coagulation	[172]
Cardiovascular	Calcific aortic valves	Tissue	iTRAQ	Tenascin-X	105 differentially expressed proteins	Blood coagulation, integrin signaling pathways, collagen fibrillogenesis	[173]
Neurological	Schizophrenia	Serum	LC-MS/MS	—	27 differentially expressed proteins	Complement and coagulation cascades, nervous system development and function, humoral immune response, inflammatory response, immune system	[174]
Neurological	Parkinson's disease (PD)	Serum	iTRAQ	Complement component 4B, apolipoprotein A-I, α-2-antiplasmin, coagulation factor V, serotransferrin, clusterin	26 differentially expressed proteins	Oxidative stress, mitochondrial dysfunction, abnormal protein aggregation, lipid metabolism, inflammation	[175]

LC-MS/MS coupled iTRAQ [183], MALDI-MS/MS in conjunction with DIGE [184], and so forth. This research has led to a commonly accepted pathway model involving glycolysis, fatty acid metabolism, and elements of the TCA and mitochondrial electron transport machinery [185]. However, proteomic heart failure biomarkers that can be used in a clinical setting remain elusive.

A recent publication by Scott et al. [186] identified six potential serum biomarkers using SELDI-TOF-MS, but none of the proposed markers were predictive for suitability of patients for clinical intervention therapy. A study by Kuznetsova et al. [187] used CE-MS to investigate urinary biomarkers in hypertension. They showed a peak pattern of 85 discriminatory peaks and, after adjustment for multiple testing, three potential biomarkers remained sufficiently different to distinguish healthy controls from patients having essential hypertension with left ventricular diastolic dysfunction. Despite these promising initial results, clinical applicability remains to be shown, since neither molecular identification of the proposed biomarkers nor consequential validation by independent techniques has been demonstrated.

9.4.3 Kidney Diseases

Glomerulonephritis is one of the major causes of renal failure and an important factor in morbidity and mortality in multifaceted kidney diseases [188]. Glomerular disease diagnosis involves assessment of renal function, quantification of proteinuria, and urine analysis by dipstick testing and microscopy. These results are useful in determining renal injury, but they are poor indicators of pre-existing or co-occurring inflammatory conditions. This drawback has led to the study of a number of estimated biomarkers, including cytokines, chemokines, and growth factors, proposed to be involved in glomerular disease. Recently, MS has been used to screen for urine biomarkers in glomerulonephritis, both in animal models and in patients at various disease stages [189]. Using SELDI-TOF-MS, Mosely et al. [190] described two urinary biomarkers with 92% sensitivity and 92% specificity that discriminate active from inactive lupus nephritis and allow detection of relapse and remission earlier than traditional clinical markers.

Similarly, via CE-MS and 2D-gel approaches, a number of urine biomarkers identified in immunoglobulin A nephropathy were put forward to discriminate between glomerulonephritis and other proteinuric glomerular diseases [191]. Interestingly, Haubitz et al. [192] observed changes in the excretion of peptide peak patterns with increasing use of an antihypertensive drug, which opens up the possibility that the observed urinary biomarkers can be used in the future to monitor the effectiveness of clinical drug treatment in glomerulonephritis. This is furthermore supported by an independent study by Rossing et al. [193], in which it was found that treatment of diabetic nephropathy patients with the angiotensin II receptor blocker candesartan significantly reduced the excretion of disease-specific biomarkers toward levels seen in healthy control urine.

Another important cause of morbidity and mortality in patients with both type 1 and type 2 diabetes mellitus is diabetic nephropathy. Microalbuminuria is an important early-onset marker of this condition, but by the time it becomes apparent, significant glomerular injury has already occurred. MS was used in a number of studies to identify urinary polypeptide patterns in patients with normoalbuminuria and less robust patterns of microalbuminuria in diabetes [193]. This suggests that it might be possible to establish a predictive model to identify patients at risk of renal injury before nephropathy becomes established.

9.4.4 Preeclampsia

Preeclampsia is the major leading cause of maternal and neonatal morbidity and mortality, accounting for 5–7% of all pregnancy-associated complications worldwide. This is a multifactorial disease that is characterized by co-occurring hypertension and proteinuria [194]. Several studies have been undertaken to understand the molecular causes that can lead to preeclampsia and understand its pathogenicity, as well as to identify potential preeclampsia biomarkers. In particular, the study undertaken by Carthy et al. [195] demonstrated with good confidence a predictive urinary biomarker panel of 50 peptides at week 28 of pregnancy prior to onset of the disease. However, it was not possible to effectively predict preeclampsia before the 28th week of pregnancy. Potential biomarkers were identified as fibrinogen α chain, collagen α chain, and uromodulin fragments.

9.4.5 Neurological Complexes

One of the most complex human biological systems, and the least well understood, is undoubtedly the brain and its peripheral pathways. A major breakthrough in 2000 by Husi et al. [196] was the discovery of a vast, tight molecular network of synaptic proteins incorporating at least 250 proteins into a substructure of post-synaptic density. Many of the proteins were known to be key players in learning and memory; however, the demonstration of a direct association and preassembly of pathways and signaling machineries to such an extent had never before been shown, neither in the context of neuroscience nor in any other system [197]. This led to the speculation that complex diseases such as Alzheimer's or Parkinson's were a result of network perturbances, where unrelated damages in genes/proteins result in the same outcome [198].

Protein biomarker discovery from serum and other biological fluids using MS-based methodologies and associated technologies for the identification of protein biomarker signatures has been applied to neuropsychiatric disorders [199] such as schizophrenia [200], neurodegenerative conditions like multiple sclerosis [201], Alzheimer's [202], Parkinson's [203], and neurodevelopmental diseases such as attention deficit hyperactivity disorder [204] and autism spectrum disorder [205]. Although the latter two conditions are less well described in proteomic pattern profiling, they are found to be associated with metabolic pattern shifts such as serotonin levels. However, on the proteomics level it can be demonstrated that latrophilin LPHN3 expression levels are strongly associated with attention deficit hyperactivity disorder [206].

In particular, a number of MS-based methods have been used to identify potential biomarkers for multiple sclerosis. Stoop et al. [207] utilized peptide and protein profiling for detection and characterization of differences between controls and multiple sclerosis patients. In a follow-up study, they used MALDI-MS for semiquantitative analysis of differentially abundant proteins [208]. A study published by Comabella et al. [209] used iTRAQ analysis of cerebrospinal fluid from multiple sclerosis patients. They found 23 proteins that were differentially expressed in individuals with worsening multiple sclerosis conditions and those remaining stable. One of the proteins differentially expressed, chitinase-3-like 1, was further validated independently with ELISA. Another study using cerebrospinal fluid from multiple sclerosis patients was reported by Mattsson et al. [210]. The researchers demonstrated a robustly altered pattern of 24 proteins, with a down-regulation of chromogranin B and secretogranin II in multiple sclerosis.

Alzheimer's disease has been the focus of a substantial number of proteomic and diagnostic biomolecule studies. For example, Korolainen et al. [211] listed 26 proteins whose expression pattern is statistically altered in this disease, and Zetterberg [212] presented 30 proteins in cerebrospinal fluid as potential Alzheimer's biomarkers. A recent study by Choi et al. [213] used an MRM-based analysis of cerebrospinal fluid samples in validation of Alzheimer's disease diagnostics. MRM quantification demonstrated a high correlation with ELISA analysis of Alzheimer's-associated biomarkers, including Aβ1-40, Aβ1-42, retinol-binding protein, and cystatin C.

Other proteomic studies used SELDI-TOF-MS to define an Alzheimer's biomarker pattern. For example, Simonsen et al. [214] generated a biomarker profile of 5 biomolecules in cerebrospinal fluid with 100% sensitivity and 97% specificity. Other approaches include Yin et al.'s combination of 1D-PAGE and 2D-PAGE with LC-MS/MS, which detected 21 proteins that showed a different abundance between this disease and controls [215]. Finehout et al. [216] used 2D-PAGE to observe 23 differences in protein expression patterns. Hu et al. [217] used 2D-DIGE to observe differences between patients with mild dementia and no cognitive disease. Finally, Abdi et al. [218] used iTRAQ to identify more than 1500 proteins and demonstrate that the expression pattern of 136 molecules is uniquely changed in cerebrospinal fluid from Alzheimer's patients compared to Parkinson's disease sufferers, patients with dementia with Lewy body and controls, and 73 proteins with a modulated expression uniquely found in Parkinson's disease patients.

A number of proteomic studies have been conducted to discover novel biomarkers for Parkinson's disease, using techniques such as iTRAQ or 2D-PAGE. One example is the DIGE approach by Maarouf et al. [219]. A recent study by Lehnert et al. [164] employed iTRAQ and MRM to demonstrate 16 candidates as discriminators of Parkinson's disease in cerebrospinal fluid, including tyrosine-kinase-non-receptor-type 13, netrin-G1, golgin-160, and apolipoprotein B-100. Pan et al. [220] used SCX in combination with LC-MS/MS and SILAC to absolute quantify five proteins associated with this disease.

9.5 PERSONALIZED MEDICINE

Tailored medical care has been fueled by the development of modern sensitive technologies to comprehensively monitor and assess DNA, RNA, protein, and metabolite data in patients. Some validated and approved molecular tests are already used in therapeutic decision making [221,222]. However, a number of methodologies require further improvements in cost-effectiveness and robustness, as well as in general applicability and response to regulatory hurdles, prior to fulfilling the promise of personalized molecular medicine. An initial problem is the ability to handle and manage the tremendous amount of data generated by high-throughput methodologies to allow further visualization and downstream data analysis, as well as data integration of multi-OMICS approaches. Additionally, computational tools that allow such multidimensional data analysis and knowledge building and extension are somewhat sparse, but they are necessary in achieving improved patient outcomes.

Systems biology, which integrates a variety of scientific disciplines with biological and medical background information, is an iterative process for describing and visualizing interconnected events within cells, tissues, and whole systems. It describes how various molecular shifts contribute to further modulated events in terms of environmental and network connectivity impacts (e.g., gene pattern shifts, modulated signaling cascades, and runaway or inhibited metabolic pathways) and ultimately altered cell behavior under disease conditions. Such a cross-disciplinary effort is necessary to understand and convert multidimensional datasets into useful information that can classify and stratify disease states and identify potential specific novel biomarkers which can be used in prognosis, diagnosis, and response to therapeutic agents. An in-depth understanding of modulated events, using targeted therapeutics on many biological and pharmacological levels, is an absolute requirement for preventing adverse outcomes and accidental side reactions, as well as for avoiding ineffective treatments on a personalized level.

The systems biology-led discovery of key elements in aberrant pathways, molecular drivers, and biomarkers has the potential to lead to novel regimes in clinical trial design, targeting subpopulations of disease cases where a clear benefit of a targeted approach has yet to be identified, and it might increase the success rate and accelerate the implementation of effective treatments in the clinical setting.

9.6 CONCLUSIONS

Pharmacogenomics, biomarker discovery, and proteomics play an important role in disease characterization and therein, therapy-appropriate samples for molecular research are extremely important.

There are several proteomic workflows available that allow for many combinations and alternatives. The main factors dictating the most appropriate proteomic workflow are the objective of the study and the type of samples available.

In proteomics, sample preparation methods are mainly aimed at reducing the complexity of biological systems and rendering samples compatible with MS analysis. MS scans provide a rich source of information containing peak patterns, distribution, and quantitation, and tandem MS/MS scans can be used to mathematically reconstruct a probable amino acid sequence using appropriate databases. Important downstream processing steps include data processing, management and visualization of results, steps that are normally achieved using more than one software tool. Although some validated and approved molecular tests are already used in therapeutic decision making, advancements in personalized molecular medicine will depend on further improvements in the cost-effectiveness and applicability of a number of technologies, the response from regulatory agencies, and the ability to handle and manage the tremendous amount of data generated by high-throughput methodologies.

ACRONYMS

2D-PAGE	2D polyacrylamide electrophoresis
APEX	absolute protein expression
CE	capillary electrophoresis
CHCA	α-cyano-4-hydroxycinnamic acid
CID	collision induced dissociation
DIGE	difference gel electrophoresis
DMF	N,N-dimethyl formamide
ECD	electron capture dissociation
emPAI	exponentially modified protein abundance index
ESI	electrospray ionization
EST	expressed sequence tag
ETD	electron transfer dissociation
FASP	filter-aided sample preparation
GO	gene ontology
HCD	high-energy collision-induced dissociation
ICAT	isotope-coded affinity tag
IEF	isoelectric focusing
IPA	ingenuity pathway analysis
LC	liquid chromatography
MALDI	matrix-assisted laser desorption-ionization
MARS	Multiple Affinity Removal System
MELDI	material-enhanced laser desorption/ionization
MiMI	Michigan Molecular Interactions
MRM	multiple reaction monitoring
MS	mass spectrometry
MudPIT	multidimensional protein identification technology
NSC	nearest shrunken centroid
PAGE	polyacrylamide gel electrophoresis
PCA	principal component analysis
pI	isoelectric point
PTM	post-translational modification
SELDI	surface-enhanced laser desorption-ionization
SILAC	stable isotope labeling by amino acids in cell culture
SPA	sinapinic acid
SRM	selected reaction monitoring
SVM	support vector machine
TOF-MS	time-of-flight mass spectrometry
TOF	time-of-flight
VIP	variable importance in the projection

REFERENCES

[1] Cox J, Mann M. Is proteomics the new genomics? Cell 2007;130(3):395–8.

[2] Tian Q, Price ND, Hood L. Systems cancer medicine: towards realization of predictive, preventive, personalized and participatory (P4) medicine. J Intern Med 2012;271(2):111–21.

[3] Balch WE, Yates III JR. Application of mass spectrometry to study proteomics and interactomics in cystic fibrosis. Methods Mol Biol 2011;742:227–47.

[4] Yates JR, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 2009;11:49–79.

[5] Sadygov RG, Yates III JR. A hypergeometric probability model for protein identification and validation using tandem mass spectral data and protein sequence databases. Anal Chem 2003;75(15):3792–8.

[6] Moore RE, Young MK, Lee TD. Method for screening peptide fragment ion mass spectra prior to database searching. J Am Soc Mass Spectrom 2000;11(5):422–6.

[7] Frewen B, MacCoss MJ. Using BiblioSpec for creating and searching tandem MS peptide libraries. Curr Protoc Bioinformatics 2007. [chapter 13: unit]

[8] Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 1999;17(10):994–9.

[9] Schnolzer M, Jedrzejewski P, Lehmann WD. Protease-catalyzed incorporation of 18O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. Electrophoresis 1996;17(5):945–53.

[10] Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem 2003;75(8):1895–904.

[11] Kumar C, Mann M. Bioinformatics analysis of mass spectrometry-based proteomics data sets. FEBS Lett 2009;583(11):1703–12.

[12] Siuti N, Kelleher NL. Decoding protein modifications using top-down mass spectrometry. Nat Methods 2007;4(10):817–21.

[13] Uttenweiler-Joseph S, Claverol S, Sylvius L, Bousquet-Dubouch MP, Burlet-Schiltz O, Monsarrat B. Toward a full characterization of the human 20S proteasome subunits and their isoforms by a combination of proteomic approaches. Methods Mol Biol 2008;484:111–30.

[14] Du Y, Parks BA, Sohn S, Kwast KE, Kelleher NL. Top-down approaches for measuring expression ratios of intact yeast proteins using Fourier transform mass spectrometry. Anal Chem 2006;78(3):686–94.

[15] Sharma S, Simpson DC, Tolic N, Jaitly N, Mayampurath AM, Smith RD, et al. Proteomic profiling of intact proteins using WAX-RPLC 2-D separations and FTICR mass spectrometry. J Proteome Res 2007;6(2):602–10.

[16] Bunger MK, Cargile BJ, Ngunjiri A, Bundy JL, Stephenson Jr JL. Automated proteomics of E. coli via top-down electron-transfer dissociation mass spectrometry. Anal Chem 2008;80(5):1459–67.

[17] Schrohl AS, Wurtz S, Kohn E, Banks RE, Nielsen HJ, Sweep FC, et al. Banking of biological fluids for studies of disease-associated protein biomarkers. Mol Cell Proteomics 2008;7(10):2061–6.

[18] Skold K, Alm H, Scholz B. The impact of biosampling procedures on molecular data interpretation. Mol Cell Proteomics 2013;12(6):1489–501.

[19] Shao C, Wang Y, Gao Y. Applications of urinary proteomics in biomarker discovery. Sci China Life Sci 2011;54(5):409–17.

[20] Fiedler GM, Ceglarek U, Leichtle A, Thiery J. Standardized preprocessing of urine for proteome analysis. Methods Mol Biol 2010;641:47–63.

[21] Thongboonkerd V, Chutipongtanate S, Kanlaya R. Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: quantity, quality, and variability. J Proteome Res 2006;5(1):183–91.

[22] Albalat A, Franke J, Gonzalez J, Mischak H, Zurbig P. Urinary proteomics based on capillary electrophoresis coupled to mass spectrometry in kidney disease. Methods Mol Biol 2013;919:203–13.

[23] McLeod WD, Nunnally HN, Cantrell PE. Dependence of EMG power spectra on electrode type. IEEE Trans Biomed Eng 1976;23(2):172–5.

[24] Luque-Garcia JL, Neubert TA. Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry. J Chromatogr A 2007;1153(1–2):259–76.

[25] Apweiler R, Aslanidis C, Deufel T, Gerstner A, Hansen J, Hochstrasser D, et al. Approaching clinical proteomics: current state and future fields of application in fluid proteomics. Clin Chem Lab Med 2009;47(6):724–44.

[26] Ndao M. Biomarker discovery in serum/plasma using surface enhanced laser desorption ionization time of flight (SELDI-TOF) mass spectrometry. Methods Mol Biol 2012;818:67–79.

[27] Darde VM, Barderas MG, Vivanco F. Depletion of high-abundance proteins in plasma by immunoaffinity subtraction for two-dimensional difference gel electrophoresis analysis. Methods Mol Biol 2007;357:351–64.

[28] Govorukhina NI, de VM, Reijmers TH, Horvatovich P, van der Zee AG, Bischoff R. Influence of clotting time on the protein composition of serum samples based on LC-MS data. J Chromatogr B Analyt Technol Biomed Life Sci 2009;877(13):1281–91.

[29] Ressom HW, Varghese RS, Abdel-Hamid M, Eissa SA, Saha D, Goldman L, et al. Analysis of mass spectral serum profiles for biomarker selection. Bioinformatics 2005;21(21):4039–45.

[30] Zhang A, Sun H, Wang P, Wang X. Salivary proteomics in biomedical research. Clin Chim Acta 2013;415:261–5.

[31] Border MB, Schwartz S, Carlson J, Dibble CF, Kohltfarber H, Offenbacher S, et al. Exploring salivary proteomes in edentulous patients with type 2 diabetes. Mol Biosyst 2012;8(4):1304–10.

[32] Henson BS, Wong DT. Collection, storage, and processing of saliva samples for downstream molecular applications. Methods Mol Biol 2010;666:21–30.

[33] Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Methods 2009;6(5):359–62.

[34] Cravatt BF, Simon GM, Yates III JR. The biological impact of mass-spectrometry-based proteomics. Nature 2007;450(7172):991–1000.

[35] Conrotto P, Souchelnytskyi S. Proteomic approaches in biological and medical sciences: principles and applications. Exp Oncol 2008;30(3):171–80.

[36] Apweiler R, Aslanidis C, Deufel T, Gerstner A, Hansen J, Hochstrasser D, et al. Approaching clinical proteomics: current state and future fields of application in cellular proteomics. Cytometry A 2009;75(10):816–32.

[37] Fournier ML, Gilmore JM, Martin-Brown SA, Washburn MP. Multidimensional separations-based shotgun proteomics. Chem Rev 2007;107(8):3654–86.

[38] Scherp P, Ku G, Coleman L, Kheterpal I. Gel-based and gel-free proteomic technologies. Methods Mol Biol 2011;702:163–90.

[39] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;227(5259):680–5.

[40] Gallagher SR. One-dimensional SDS Gel electrophoresis of proteins. Curr Protoc Protein Sci 2012. [chapter 10: unit-44]

[41] Yokono T, Mineki R, Taka H, Kotaniguchi H, Murayama K. Improvement of automatic in-gel digestion by in situ alkylation of proteins. J Biomol Tech 2003;14(3):191–6.

[42] O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. J Biol Chem 1975;250(10):4007–21.

[43] Fichmann J, Westermeier R. 2-D protein gel electrophoresis. An overview. Methods Mol Biol 1999;112:1–7.

[44] Weiss W, Gorg A. High-resolution two-dimensional electrophoresis. Methods Mol Biol 2009;564:13–32.

[45] Dunn MJ. Two-dimensional polyacrylamide gel electrophoresis for cardiovascular proteomics. Methods Mol Biol 2007;357:3–13.

[46] Lauber WM, Carroll JA, Dufield DR, Kiesel JR, Radabaugh MR, Malone JP. Mass spectrometry compatibility of two-dimensional gel protein stains. Electrophoresis 2001;22(5):906–18.

[47] Lilley KS, Friedman DB. All about DIGE: quantification technology for differential-display 2D-gel proteomics. Expert Rev Proteomics 2004;1(4):401–9.

[48] Martinez-Pinna R, Lopez JA, Ramos-Mozo P, Blanco-Colio LM, Camafeita E, Calvo E, et al. Identification of novel biomarkers of abdominal aortic aneurysms by 2D-DIGE and MALDI-MS from AAA-thrombus-conditioned media. Methods Mol Biol 2013;1000:91–101.

[49] Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 1996;68(5):850–8.

[50] Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, et al. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc Natl Acad Sci U S A 1996;93(25):14440–5.

[51] Zhang Q, Menon R, Deutsch EW, Pitteri SJ, Faca VM, Wang H, et al. A mouse plasma peptide atlas as a resource for disease proteomics. Genome Biol 2008;9(6):R93.

[52] Wang H, Hanash S. Intact-protein analysis system for discovery of serum-based disease biomarkers. Methods Mol Biol 2011;728:69–85.

[53] Jorgenson JW, Lukacs KD. Free-zone electrophoresis in glass capillaries. Clin Chem 1981;27(9):1551–3.

[54] Sarker SD, Nahar L. Hyphenated techniques and their applications in natural products analysis. Methods Mol Biol 2012;864:301–40.

[55] Loo JA, Udseth HR, Smith RD. Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry. Anal Biochem 1989;179(2):404–12.

[56] Mischak H, Schanstra JP. CE-MS in biomarker discovery, validation, and clinical application. Proteomics Clin Appl 2011;5(1–2):9–23.

[57] Albalat A, Bitsika V, Zurbig P, Siwy J, Mullen W. High-resolution proteome/peptidome analysis of body fluids by capillary electrophoresis coupled with MS. Methods Mol Biol 2013;984:153–65.

[58] Haselberg R, de Jong GJ, Somsen GW. CE-MS for the analysis of intact proteins 2010–2012. Electrophoresis 2013;34(1):99–112.

[59] Capriotti AL, Cavaliere C, Foglia P, Samperi R, Lagana A. Intact protein separation by chromatographic and/or electrophoretic techniques for top-down proteomics. J Chromatogr A 2011;1218(49):8760–76.

[60] Florens L, Washburn MP. Proteomic analysis by multidimensional protein identification technology. Methods Mol Biol 2006;328:159–75.

[61] Di SD, Brambilla F, Mauri PL. Multidimensional protein identification technology for direct-tissue proteomics of heart. Methods Mol Biol 2013;1005:25–38.

[62] Chapman K. The ProteinChip Biomarker System from Ciphergen Biosystems: a novel proteomics platform for rapid biomarker discovery and validation. Biochem Soc Trans 2002;30(2):82–7.

[63] Langbein S. Identification of disease biomarkers by profiling of serum proteins using SELDI-TOF mass spectrometry. Methods Mol Biol 2008;439:191–7.

[64] Peschke M, Verkerk UH, Kebarle P. Prediction of the charge states of folded proteins in electrospray ionization. Eur J Mass Spectrom (Chichester, Eng) 2004;10(6):993–1002.

[65] Shen Y, Moore RJ, Zhao R, Blonder J, Auberry DL, Masselon C, et al. High-efficiency on-line solid-phase extraction coupling to 15–150-microm-i.d. column liquid chromatography for proteomic analysis. Anal Chem 2003;75(14):3596–605.

[66] Wang H, Hanash S. Electrospray mass spectrometry for quantitative plasma proteome analysis. Methods Mol Biol 2009;564:227–42.

[67] Silvertand LH, Torano JS, de Jong GJ, van Bennekom WP. Improved repeatability and matrix-assisted desorption/ionization—time of flight mass spectrometry compatibility in capillary isoelectric focusing. Electrophoresis 2008;29(10):1985–96.

[68] Chen Y, Wall D, Lubman DM. Rapid identification and screening of proteins from whole cell lysates of human erythroleukemia cells in the liquid phase, using non-porous reversed phase high-performance liquid chromatography separations of proteins followed by matrix-assisted [correction of multi-assisted] laser desorption/ionization mass spectrometry analysis and sequence database searching. Rapid Commun Mass Spectrom 1998; 12(24):1994–2003.

[69] Shen Y, Zhao R, Belov ME, Conrads TP, Anderson GA, Tang K, et al. Packed capillary reversed-phase liquid chromatography with high-performance electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for proteomics. Anal Chem 2001;73(8):1766–75.

[70] Laiko VV, Baldwin MA, Burlingame AL. Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry. Anal Chem 2000;72(4):652–7.

[71] Merchant M, Weinberger SR. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis 2000;21(6):1164–77.

[72] Feuerstein I, Najam-ul-Haq M, Rainer M, Trojer L, Bakry R, Aprilita NH, et al. Material-enhanced laser desorption/ionization (MELDI)—a new protein profiling tool utilizing specific carrier materials for time of flight mass spectrometric analysis. J Am Soc Mass Spectrom 2006;17(9):1203–8.

[73] Hughes C, Ma B, Lajoie GA. De novo sequencing methods in proteomics. Methods Mol Biol 2010;604:105–21.

[74] Lin D, Tabb DL, Yates III JR. Large-scale protein identification using mass spectrometry. Biochim Biophys Acta 2003;1646(1–2):1–10.

[75] MacCoss MJ. Computational analysis of shotgun proteomics data. Curr Opin Chem Biol 2005;9(1):88–94.

[76] Johnson RS, Davis MT, Taylor JA, Patterson SD. Informatics for protein identification by mass spectrometry. Methods 2005;35(3):223–36.

[77] Pedrioli PG, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Raught B, et al. A common open representation of mass spectrometry data and its application to proteomics research. Nat Biotechnol 2004;22(11):1459–66.

[78] Martens L, Chambers M, Sturm M, Kessner D, Levander F, Shofstahl J, et al. MzML—a community standard for mass spectrometry data. Mol Cell Proteomics 2011;10(1):R110.

[79] Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, et al. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. Rapid Commun Mass Spectrom 2003;17(20):2337–42.

[80] Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 1999;20(18):3551–67.

[81] Collins MO, Yu L, Husi H, Blackstock WP, Choudhary JS, Grant SG. Robust enrichment of phosphorylated species in complex mixtures by sequential protein and peptide metal-affinity chromatography and analysis by tandem mass spectrometry. Sci STKE 2005;2005(298):l6.

[82] Schmidt A, Kellermann J, Lottspeich F. A novel strategy for quantitative proteomics using isotope-coded protein labels. Proteomics 2005;5(1):4–15.

[83] Yi EC, Li XJ, Cooke K, Lee H, Raught B, Page A, et al. Increased quantitative proteome coverage with (13)C/(12)C-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. Proteomics 2005;5(2):380–7.

[84] Wang M, You J. Mass spectrometry for protein quantification in biomarker discovery. Methods Mol Biol 2012;815:199–225.

[85] Calvo E, Martinez-Pinna R, Ramos-Mozo P, Pastor-Vargas C, Camafeita E, Egido J, et al. Unraveling biomarkers of abdominal aortic aneurisms by iTRAQ analysis of depleted plasma. Methods Mol Biol 2013;1000:157–66.

[86] Onsongo G, Stone MD, Van Riper SK, Chilton J, Wu B, Higgins L, et al. LTQ-iQuant: a freely available software pipeline for automated and accurate protein quantification of isobaric tagged peptide data from LTQ instruments. Proteomics 2010;10(19):3533–8.

[87] Arntzen MO, Koehler CJ, Barsnes H, Berven FS, Treumann A, Thiede B. IsobariQ: software for isobaric quantitative proteomics using IPTL, iTRAQ, and TMT. J Proteome Res 2011;10(2):913–20.

[88] Wang P, Yang P, Yang JY. OCAP: an open comprehensive analysis pipeline for iTRAQ. Bioinformatics 2012;28(10):1404–5.

[89] Zou X, Zhao M, Shen H, Zhao X, Tong Y, Wang Q, et al. MilQuant: a free, generic software tool for isobaric tagging-based quantitation. J Proteomics 2012;75(17):5516–22.

[90] Eisenacher M, Kohl M, Wiese S, Hebeler R, Meyer HE, Warscheid B, et al. Find pairs: the module for protein quantification of the PeakQuant software suite. OMICS 2012;16(9):457–67.

[91] Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci U S A 1999;96(12):6591–6.

[92] Sury MD, Chen JX, Selbach M. The SILAC fly allows for accurate protein quantification *in vivo*. Mol Cell Proteomics 2010;9(10):2173–83.

[93] Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 2002;1(5):376–86.

[94] Liao Z, Wan Y, Thomas SN, Yang AJ. IsoQuant: a software tool for stable isotope labeling by amino acids in cell culture-based mass spectrometry quantitation. Anal Chem 2012;84(10):4535–43.

[95] Hoopmann MR, Chavez JD, Bruce JE. SILACtor: software to enable dynamic SILAC studies. Anal Chem 2011;83(22):8403–10.

[96] Kuhn E, Wu J, Karl J, Liao H, Zolg W, Guild B. Quantification of C-reactive protein in the serum of patients with rheumatoid

arthritis using multiple reaction monitoring mass spectrometry and 13C-labeled peptide standards. Proteomics 2004;4(4):1175–86.

[97] Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol 2010;2010:840518.

[98] Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 2005;4(10):1487–502.

[99] Wong JW, Cagney G. An overview of label-free quantitation methods in proteomics by mass spectrometry. Methods Mol Biol 2010;604:273–83.

[100] Liu H, Sadygov RG, Yates III JR. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 2004;76(14):4193–201.

[101] Florens L, Carozza MJ, Swanson SK, Fournier M, Coleman MK, Workman JL, et al. Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods 2006;40(4):303–11.

[102] Braisted JC, Kuntumalla S, Vogel C, Marcotte EM, Rodrigues AR, Wang R, et al. The APEX Quantitative Proteomics Tool: generating protein quantitation estimates from LC-MS/MS proteomics results. BMC Bioinformatics 2008;9:529.

[103] Kiyonami R, Domon B. Selected reaction monitoring applied to quantitative proteomics. Methods Mol Biol 2010;658:155–66.

[104] Bessant C, Fan J. Computational approaches to selected reaction monitoring assay design. Methods Mol Biol 2013;1007:219–35.

[105] Krokhin OV, Craig R, Spicer V, Ens W, Standing KG, Beavis RC, et al. An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. Mol Cell Proteomics 2004;3(9):908–19.

[106] Schmidt C, Urlaub H. Absolute quantification of proteins using standard peptides and multiple reaction monitoring. Methods Mol Biol 2012;893:249–65.

[107] Colangelo CM, Chung L, Bruce C, Cheung KH. Review of software tools for design and analysis of large scale MRM proteomic datasets. Methods 2013;61(3):287–98.

[108] Perez-Riverol Y, Wang R, Hermjakob H, Muller M, Vesada V, Vizcaino JA. Open source libraries and frameworks for mass spectrometry based proteomics: a developer's perspective. Biochim Biophys Acta 2014;1844 (1 Pt A):63–76.

[109] Levner I. Feature selection and nearest centroid classification for protein mass spectrometry. BMC Bioinformatics 2005;6:68.

[110] Liggett WS, Barker PE, Semmes OJ, Cazares LH. Measurement reproducibility in the early stages of biomarker development. Dis Markers 2004;20(6):295–307.

[111] Guo Y, Balasubramanian R. Comparative evaluation of classifiers in the presence of statistical interactions between features in high dimensional data settings. Int J Biostat 2012;8(1).

[112] Yates III JR, Eng JK, McCormack AL, Schieltz D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal Chem 1995;67(8):1426–36.

[113] Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de BB, et al. Reactome a knowledgebase of biological pathways. Nucleic Acids Res 2005;33(Database issue):D428–32.

[114] van Iersel MP, Kelder T, Pico AR, Hanspers K, Coort S, Conklin BR, et al. Presenting and exploring biological pathways with PathVisio. BMC Bioinformatics 2008;9:399.

[115] Yeung N, Cline MS, Kuchinsky A, Smoot ME, Bader GD. Exploring biological networks with Cytoscape software. Curr Protoc Bioinformatics 2008. [chapter 8: unit]

[116] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215(3):403–10.

[117] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22(22):4673–80.

[118] Husi H, Skipworth RJ, Fearon KC, Ross JA. LSCluster, a large-scale sequence clustering and aligning software for use in partial identity mapping and splice-variant analysis. J Proteomics 2013;84:185–9.

[119] Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009;25(8):1091–3.

[120] Gao J, Ade AS, Tarcea VG, Weymouth TE, Mirel BR, Jagadish HV, et al. Integrating and annotating the interactome using the MiMI plugin for cytoscape. Bioinformatics 2009;25(1):137–8.

[121] Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. Bioinformatics 2010;26(22):2927–8.

[122] Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. Bioinformatics 2008;24(2):282–4.

[123] Xia T, Hemert JV, Dickerson JA. OmicsAnalyzer: a Cytoscape plug-in suite for modeling omics data. Bioinformatics 2010;26(23):2995–6.

[124] Webb-Robertson BJ, Matzke MM, Jacobs JM, Pounds JG, Waters KM. A statistical selection strategy for normalization procedures in LC-MS proteomics experiments through dataset-dependent ranking of normalization scaling factors. Proteomics 2011;11(24):4736–41.

[125] Smit S, Hoefsloot HC, Smilde AK. Statistical data processing in clinical proteomics. J Chromatogr B Analyt Technol Biomed Life Sci 2008;866(1–2):77–88.

[126] Cherkassky V, Ma Y. Another look at statistical learning theory and regularization. Neural Netw 2009;22(7):958–69.

[127] Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002;99(10):6567–72.

[128] Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 2004;573(1–3):83–92.

[129] Smit S, van Breemen MJ, Hoefsloot HC, Smilde AK, Aerts JM, de Koster CG. Assessing the statistical validity of proteomics based biomarkers. Anal Chim Acta 2007;592(2):210–7.

[130] Bijlsma S, Bobeldijk I, Verheij ER, Ramaker R, Kochhar S, Macdonald IA, et al. Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. Anal Chem 2006;78(2):567–74.

[131] Yang J, Zhao X, Liu X, Wang C, Gao P, Wang J, et al. High performance liquid chromatography-mass spectrometry for metabonomics: potential biomarkers for acute deterioration of liver function in chronic hepatitis B. J Proteome Res 2006;5(3):554–61.

[132] Shen C, Breen TE, Dobrolecki LE, Schmidt CM, Sledge GW, Miller KD, et al. Comparison of computational algorithms for the classification of liver cancer using SELDI mass spectrometry: a case study. Cancer Inform 2007;3:329–39.

[133] Titulaer MK, Siccama I, Dekker LJ, van Rijswijk AL, Heeren RM, Sillevis Smitt PA, et al. A database application for pre-processing, storage and comparison of mass spectra derived from patients and controls. BMC Bioinformatics 2006;7:403.

[134] Mancuso F, Bunkenborg J, Wierer M, Molina H. Data extraction from proteomics raw data: an evaluation of nine tandem MS tools using a large Orbitrap data set. J Proteomics 2012;75(17):5293–303.

[135] Kandasamy K, Pandey A, Molina H. Evaluation of several MS/MS search algorithms for analysis of spectra derived from electron transfer dissociation experiments. Anal Chem 2009;81(17):7170–80.

[136] Wenger CD, Coon JJ. A proteomics search algorithm specifically designed for high-resolution tandem mass spectra. J Proteome Res 2013;12(3):1377–86.

[137] Cox J, Matic I, Hilger M, Nagaraj N, Selbach M, Olsen JV, et al. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nat Protoc 2009;4(5):698–705.

[138] Veenstra TD, Prieto DA, Conrads TP. Proteomic patterns for early cancer detection. Drug Discov Today 2004;9(20):889–97.

[139] Husi H, Stephens N, Cronshaw A, MacDonald A, Gallagher I, Greig C, et al. Proteomic analysis of urinary upper gastrointestinal cancer markers. Proteomics Clin Appl 2011;5(5–6):289–99.

[140] Peres S, Molina L, Salvetat N, Granier C, Molina F. A new method for 2D gel spot alignment: application to the analysis of large sample sets in clinical proteomics. BMC Bioinformatics 2008;9:460.

[141] Ballardini R, Benevento M, Arrigoni G, Pattini L, Roda A. MassUntangler: a novel alignment tool for label-free liquid chromatography-mass spectrometry proteomic data. J Chromatogr A 2011;1218(49):8859–68.

[142] Zhang W, Li F, Nie L. Integrating multiple 'omics' analysis for microbial biology: application and methodologies. Microbiology 2010;156(Pt 2):287–301.

[143] Jorgensen C, Locard-Paulet M. Analysing signalling networks by mass spectrometry. Amino Acids 2012;43(3):1061–74.

[144] Dumas ME. Metabolome 2.0: quantitative genetics and network biology of metabolic phenotypes. Mol Biosyst 2012;8(10):2494–502.

[145] Kohl M, Megger DA, Trippler M, Meckel H, Ahrens M, Bracht T, et al. A practical data processing workflow for multi-OMICS projects. Biochim Biophys Acta 2014;1844(1 Pt A):52–62.

[146] Hood L, Balling R, Auffray C. Revolutionizing medicine in the 21st century through systems approaches. Biotechnol J 2012;7(8):992–1001.

[147] Waters KM, Pounds JG, Thrall BD. Data merging for integrated microarray and proteomic analysis. Brief Funct Genomic Proteomic 2006;5(4):261–72.

[148] Sefcik LS, Wilson JL, Papin JA, Botchwey EA. Harnessing systems biology approaches to engineer functional microvascular networks. Tissue Eng Part B Rev 2010;16(3):361–70.

[149] Katoh M, Katoh M. Bioinformatics for cancer management in the post-genome era. Technol Cancer Res Treat 2006;5(2):169–75.

[150] Quackenbush J. Data standards for 'omic' science. Nat Biotechnol 2004;22(5):613–4.

[151] Roxo-Rosa M, da CG, Luider TM, Scholte BJ, Coelho AV, Amaral MD, et al. Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006;6(7):2314–25.

[152] Zhou G, Li H, DeCamp D, Chen S, Shu H, Gong Y, et al. 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. Mol Cell Proteomics 2002;1(2):117–24.

[153] Steel LF, Shumpert D, Trotter M, Seeholzer SH, Evans AA, London WT, et al. A strategy for the comparative analysis of serum proteomes for the discovery of biomarkers for hepatocellular carcinoma. Proteomics 2003;3(5):601–9.

[154] Lantz RC, Lynch BJ, Boitano S, Poplin GS, Littau S, Tsaprailis G, et al. Pulmonary biomarkers based on alterations in protein expression after exposure to arsenic. Environ Health Perspect 2007;115(4):586–91.

[155] Mauri P, Scarpa A, Nascimbeni AC, Benazzi L, Parmagnani E, Mafficini A, et al. Identification of proteins released by pancreatic cancer cells by multidimensional protein identification technology: a strategy for identification of novel cancer markers. FASEB J 2005;19(9):1125–7.

[156] Watson CJ, Ledwidge MT, Phelan D, Collier P, Byrne JC, Dunn MJ, et al. Proteomic analysis of coronary sinus serum reveals leucine-rich alpha2-glycoprotein as a novel biomarker of ventricular dysfunction and heart failure. Circ Heart Fail 2011;4(2):188–97.

[157] de la Cuesta F, Barderas MG, Calvo E, Zubiri I, Maroto AS, Darde VM, et al. Secretome analysis of atherosclerotic and non-atherosclerotic arteries reveals dynamic extracellular remodeling during pathogenesis. J Proteomics 2012;75(10):2960–71.

[158] Wang Q, Li D, Zhang W, Tang B, Li QQ, Li L. Evaluation of proteomics-identified CCL18 and CXCL1 as circulating tumor markers for differential diagnosis between ovarian carcinomas and benign pelvic masses. Int J Biol Markers 2011;26(4):262–73.

[159] Liu C, Pan C, Liang Y. Screening and identification of serum proteomic biomarkers for gastric adenocarcinoma. Exp Ther Med 2012;3(6):1005–9.

[160] Delles C, Schiffer E, von Zur MC, Peter K, Rossing P, Parving HH, et al. Urinary proteomic diagnosis of coronary artery disease: identification and clinical validation in 623 individuals. J Hypertens 2010;28(11):2316–22.

[161] Jahn H, Wittke S, Zurbig P, Raedler TJ, Arlt S, Kellmann M, et al. Peptide fingerprinting of Alzheimer's disease in cerebrospinal fluid: identification and prospective evaluation of new synaptic biomarkers. PLoS One 2011;6(10):e26540.

[162] Kroksveen AC, Aasebo E, Vethe H, Van PV, Franciotta D, Teunissen CE, et al. Discovery and initial verification of differentially abundant proteins between multiple sclerosis patients and controls using iTRAQ and SID-SRM. J Proteomics 2013;78:312–25.

[163] Gautam P, Nair SC, Gupta MK, Sharma R, Polisetty RV, Uppin MS, et al. Proteins with altered levels in plasma from glioblastoma patients as revealed by iTRAQ-based quantitative proteomic analysis. PLoS One 2012;7(9):e46153.

[164] Lehnert S, Jesse S, Rist W, Steinacker P, Soininen H, Herukka SK, et al. ITRAQ and multiple reaction monitoring as proteomic tools for biomarker search in cerebrospinal fluid of patients with Parkinson's disease dementia. Exp Neurol 2012;234(2):499–505.

[165] Liu H, Ong SE, Badu-Nkansah K, Schindler J, White FM, Hynes RO. CUB-domain-containing protein 1 (CDCP1) activates Src to promote melanoma metastasis. Proc Natl Acad Sci U S A 2011;108(4):1379–84.

[166] Park J, Cha DH, Lee SJ, Kim YN, Kim YH, Kim KP. Discovery of the serum biomarker proteins in severe preeclampsia by proteomic analysis. Exp Mol Med 2011;43(7):427–35.

[167] Zhao Y, Jia W, Sun W, Jin W, Guo L, Wei J, et al. Combination of improved (18)O incorporation and multiple reaction

monitoring: a universal strategy for absolute quantitative verification of serum candidate biomarkers of liver cancer. J Proteome Res 2010;9(6):3319–27.

[168] Plymoth A, Hainaut P. Proteomics beyond proteomics: toward clinical applications. Curr Opin Oncol 2011;23(1):77–82.

[169] Ummanni R, Mundt F, Pospisil H, Venz S, Scharf C, Barett C, et al. Identification of clinically relevant protein targets in prostate cancer with 2D-DIGE coupled mass spectrometry and systems biology network platform. PLoS One 2011;6(2):e16833.

[170] Sodek KL, Evangelou AI, Ignatchenko A, Agochiya M, Brown TJ, Ringuette MJ, et al. Identification of pathways associated with invasive behavior by ovarian cancer cells using multidimensional protein identification technology (MudPIT). Mol Biosyst 2008;4(7):762–73.

[171] Gollapalli K, Ray S, Srivastava R, Renu D, Singh P, Dhali S, et al. Investigation of serum proteome alterations in human glioblastoma multiforme. Proteomics 2012;12(14):2378–90.

[172] Kaur P, Rizk NM, Ibrahim S, Younes N, Uppal A, Dennis K, et al. ITRAQ-based quantitative protein expression profiling and MRM verification of markers in type 2 diabetes. J Proteome Res 2012;11(11):5527–39.

[173] Matsumoto K, Satoh K, Maniwa T, Araki A, Maruyama R, Oda T. Noticeable decreased expression of tenascin-X in calcific aortic valves. Connect Tissue Res 2012;53(6):460–8.

[174] Li Y, Zhou K, Zhang Z, Sun L, Yang J, Zhang M, et al. Label-free quantitative proteomic analysis reveals dysfunction of complement pathway in peripheral blood of schizophrenia patients: evidence for the immune hypothesis of schizophrenia. Mol Biosyst 2012;8(10):2664–71.

[175] Zhang X, Yin X, Yu H, Liu X, Yang F, Yao J, et al. Quantitative proteomic analysis of serum proteins in patients with Parkinson's disease using an isobaric tag for relative and absolute quantification labeling, two-dimensional liquid chromatography, and tandem mass spectrometry. Analyst 2012;137(2):490–5.

[176] Verma M, Wright Jr GL, Hanash SM, Gopal-Srivastava R, Srivastava S. Proteomic approaches within the NCI early detection research network for the discovery and identification of cancer biomarkers. Ann NY Acad Sci 2001;945:103–15.

[177] Tsui KH, Tang P, Lin CY, Chang PL, Chang CH, Yung BY. Bikunin loss in urine as useful marker for bladder carcinoma. J Urol 2010;183(1):339–44.

[178] Chen YT, Chen CL, Chen HW, Chung T, Wu CC, Chen CD, et al. Discovery of novel bladder cancer biomarkers by comparative urine proteomics using iTRAQ technology. J Proteome Res 2010;9(11):5803–15.

[179] Tan LB, Chen KT, Yuan YC, Liao PC, Guo HR. Identification of urine PLK2 as a marker of bladder tumors by proteomic analysis. World J Urol 2010;28(1):117–22.

[180] Petri AL, Simonsen AH, Yip TT, Hogdall E, Fung ET, Lundvall L, et al. Three new potential ovarian cancer biomarkers detected in human urine with equalizer bead technology. Acta Obstet Gynecol Scand 2009;88(1):18–26.

[181] Metzger J, Negm AA, Plentz RR, Weismuller TJ, Wedemeyer J, Karlsen TH, et al. Urine proteomic analysis differentiates cholangiocarcinoma from primary sclerosing cholangitis and other benign biliary disorders. Gut 2013;62(1):122–30.

[182] Dubois E, Fertin M, Burdese J, Amouyel P, Bauters C, Pinet F. Cardiovascular proteomics: translational studies to develop novel biomarkers in heart failure and left ventricular remodeling. Proteomics Clin Appl 2011;5(1–2):57–66.

[183] Jullig M, Hickey AJ, Chai CC, Skea GL, Middleditch MJ, Costa S, et al. Is the failing heart out of fuel or a worn engine running rich? A study of mitochondria in old spontaneously hypertensive rats. Proteomics 2008;8(12):2556–72.

[184] Kotlo K, Johnson KR, Grillon JM, Geenen DL, deTombe P, Danziger RS. Phosphoprotein abundance changes in hypertensive cardiac remodeling. J Proteomics 2012;77:1–13.

[185] Meng C, Jin X, Xia L, Shen SM, Wang XL, Cai J, et al. Alterations of mitochondrial enzymes contribute to cardiac hypertrophy before hypertension development in spontaneously hypertensive rats. J Proteome Res 2009;8(5):2463–75.

[186] Scott PA, Zeidan B, Ng LL, Zeb M, Rosengarten JA, Garbis S, et al. Proteomic profiling to identify prognostic biomarkers in heart failure. *In Vivo* 2012 November;26(6):875–82.

[187] Kuznetsova T, Mischak H, Mullen W, Staessen JA. Urinary proteome analysis in hypertensive patients with left ventricular diastolic dysfunction. Eur Heart J 2012;33(18):2342–50.

[188] Barratt J, Topham P. Urine proteomics: the present and future of measuring urinary protein components in disease. CMAJ 2007;177(4):361–8.

[189] Boumpas DT, Kuroiwa T. Urine proteomics for the early diagnosis of acute glomerulonephritis: has the time come? Arthritis Rheum 2007;56(3):705–9.

[190] Mosley K, Tam FW, Edwards RJ, Crozier J, Pusey CD, Lightstone L. Urinary proteomic profiles distinguish between active and inactive lupus nephritis. Rheumatology (Oxford) 2006;45(12):1497–504.

[191] Park MR, Wang EH, Jin DC, Cha JH, Lee KH, Yang CW, et al. Establishment of a 2-D human urinary proteomic map in IgA nephropathy. Proteomics 2006;6(3):1066–76.

[192] Haubitz M, Wittke S, Weissinger EM, Walden M, Rupprecht HD, Floege J, et al. Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy. Kidney Int 2005;67(6):2313–20.

[193] Rossing K, Mischak H, Parving HH, Christensen PK, Walden M, Hillmann M, et al. Impact of diabetic nephropathy and angiotensin II receptor blockade on urinary polypeptide patterns. Kidney Int 2005;68(1):193–205.

[194] Mary S, Patil GV, Kulkarni AV, Kulkarni MJ, Joshi SR, Mehendale SS, et al. Dynamic proteome in enigmatic preeclampsia: an account of molecular mechanisms and biomarker discovery. Proteomics Clin Appl 2012;6(1–2):79–90.

[195] Carty DM, Siwy J, Brennand JE, Zurbig P, Mullen W, Franke J, et al. Urinary proteomics for prediction of preeclampsia. Hypertension 2011;57(3):561–9.

[196] Husi H, Ward MA, Choudhary JS, Blackstock WP, Grant SG. Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. Nat Neurosci 2000;3(7):661–9.

[197] Husi H, Grant SG. Proteomics of the nervous system. Trends Neurosci 2001;24(5):259–66.

[198] Husi H. NMDA receptors, neural pathways, and protein interaction databases. Int Rev Neurobiol 2004;61:49–77.

[199] Kroksveen AC, Opsahl JA, Aye TT, Ulvik RJ, Berven FS. Proteomics of human cerebrospinal fluid: discovery and verification of biomarker candidates in neurodegenerative diseases using quantitative proteomics. J Proteomics 2011;74(4):371–88.

[200] Martins-de-Souza D, Guest PC, Rahmoune H, Bahn S. Proteomic approaches to unravel the complexity of schizophrenia. Expert Rev Proteomics 2012;9(1):97–108.

[201] Jia Y, Wu T, Jelinek CA, Bielekova B, Chang L, Newsome S, et al. Development of protein biomarkers in cerebrospinal fluid

for secondary progressive multiple sclerosis using selected reaction monitoring mass spectrometry (SRM-MS). Clin Proteomics 2012;9(1):9.

[202] Perrin RJ, Payton JE, Malone JP, Gilmore P, Davis AE, Xiong C, et al. Quantitative label-free proteomics for discovery of biomarkers in cerebrospinal fluid: assessment of technical and inter-individual variation. PLoS One 2013;8(5):e64314.

[203] Constantinescu R, Mondello S. Cerebrospinal fluid biomarker candidates for parkinsonian disorders. Front Neurol 2012;3:187.

[204] Moriarty M, Lee A, O'Connell B, Kelleher A, Keeley H, Furey A. Development of an LC-MS/MS method for the analysis of serotonin and related compounds in urine and the identification of a potential biomarker for attention deficit hyperactivity/hyperkinetic disorder. Anal Bioanal Chem 2011;401(8):2481–93.

[205] Taurines R, Dudley E, Conner AC, Grassl J, Jans T, Guderian F, et al. Serum protein profiling and proteomics in autistic spectrum disorder using magnetic bead-assisted mass spectrometry. Eur Arch Psychiatry Clin Neurosci 2010;260(3):249–55.

[206] O'Sullivan ML, de Wit J, Savas JN, Comoletti D, Otto-Hitt S, Yates III JR, et al. FLRT proteins are endogenous latrophilin ligands and regulate excitatory synapse development. Neuron 2012;73(5):903–10.

[207] Stoop MP, Dekker LJ, Titulaer MK, Burgers PC, Sillevissmitt PA, Luider TM, et al. Multiple sclerosis-related proteins identified in cerebrospinal fluid by advanced mass spectrometry. Proteomics 2008;8(8):1576–85.

[208] Stoop MP, Dekker LJ, Titulaer MK, Lamers RJ, Burgers PC, Sillevis Smitt PA, et al. Quantitative matrix-assisted laser desorption ionization-fourier transform ion cyclotron resonance (MALDI-FT-ICR) peptide profiling and identification of multiple-sclerosis-related proteins. J Proteome Res 2009;8(3):1404–14.

[209] Comabella M, Fernandez M, Martin R, Rivera-Valve S, Borras E, Chiva C, et al. Cerebrospinal fluid chitinase 3-like 1 levels are associated with conversion to multiple sclerosis. Brain 2010;133(Pt 4):1082–93.

[210] Mattsson N, Ruetschi U, Podust VN, Stridsberg M, Li S, Andersen O, et al. Cerebrospinal fluid concentrations of peptides derived from chromogranin B and secretogranin II are decreased in multiple sclerosis. J Neurochem 2007;103(5):1932–9.

[211] Korolainen MA, Nyman TA, Aittokallio T, Pirttila T. An update on clinical proteomics in Alzheimer's research. J Neurochem 2010;112(6):1386–414.

[212] Zetterberg H, Ruetschi U, Portelius E, Brinkmalm G, Andreasson U, Blennow K, et al. Clinical proteomics in neurodegenerative disorders. Acta Neurol Scand 2008;118(1):1–11.

[213] Choi YS, Hou S, Choe LH, Lee KH. Targeted human cerebrospinal fluid proteomics for the validation of multiple Alzheimer's disease biomarker candidates. J Chromatogr B Analyt Technol Biomed Life Sci 2013;930:129–35.

[214] Simonsen AH, McGuire J, Podust VN, Davies H, Minthon L, Skoog I, et al. Identification of a novel panel of cerebrospinal fluid biomarkers for Alzheimer's disease. Neurobiol Aging 2008;29(7):961–8.

[215] Yin GN, Lee HW, Cho JY, Suk K. Neuronal pentraxin receptor in cerebrospinal fluid as a potential biomarker for neurodegenerative diseases. Brain Res 2009;1265:158–70.

[216] Finehout EJ, Franck Z, Choe LH, Relkin N, Lee KH. Cerebrospinal fluid proteomic biomarkers for Alzheimer's disease. Ann Neurol 2007;61(2):120–9.

[217] Hu Y, Malone JP, Fagan AM, Townsend RR, Holtzman DM. Comparative proteomic analysis of intra- and interindividual variation in human cerebrospinal fluid. Mol Cell Proteomics 2005;4(12):2000–9.

[218] Abdi F, Quinn JF, Jankovic J, McIntosh M, Leverenz JB, Peskind E, et al. Detection of biomarkers with a multiplex quantitative proteomic platform in cerebrospinal fluid of patients with neurodegenerative disorders. J Alzheimers Dis 2006;9(3):293–348.

[219] Maarouf CL, Beach TG, Adler CH, Shill HA, Sabbagh MN, Wu T, et al. Cerebrospinal fluid biomarkers of neuropathologically diagnosed Parkinson's disease subjects. Neurol Res 2012;34(7):669–76.

[220] Pan S, Rush J, Peskind ER, Galasko D, Chung K, Quinn J, et al. Application of targeted quantitative proteomics analysis in human cerebrospinal fluid using a liquid chromatography matrix-assisted laser desorption/ionization time-of-flight tandem mass spectrometer (LC MALDI TOF/TOF) platform. J Proteome Res 2008;7(2):720–30.

[221] Milward EA, Daneshi N, Johnstone DM. Emerging real-time technologies in molecular medicine and the evolution of integrated 'pharmacomics' approaches to personalized medicine and drug discovery. Pharmacol Ther 2012;136(3):295–304.

[222] Honda K, Ono M, Shitashige M, Masuda M, Kamita M, Miura N, et al. Proteomic approaches to the discovery of cancer biomarkers for early detection and personalized medicine. Jpn J Clin Oncol 2013;43(2):103–9.

This page is intentionally left blank

Chapter 10

Metabolomics

Karl Burgess,[a,b] **Naomi Rankin,**[a,c] **and Stefan Weidt**[a,b]

[a]*Glasgow Polyomics, University of Glasgow, UK,* [b]*Institute of Infection, Immunity and Inflammation,*
University of Glasgow, UK, [c]*Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK*

Chapter Outline

10.1 Overview 181
  10.1.1 The Metabolome and Metabolomics 181
  10.1.2 Biochemistry of the Metabolome 182
  10.1.3 Applications of Metabolomics in Clinical and Biomedical Research 183
  10.1.4 Experimental Design for Metabolomics 183
10.2 Metabolomics techniques and technologies 184
  10.2.1 Fingerprinting 184
  10.2.2 Metabolite Pool Analysis 185
  10.2.3 Metabolite Profiling by MS 185
  10.2.4 Metabolite Profiling Using NMR 185
  10.2.5 Flux Analysis 187
10.3 Analytical platform 187
  10.3.1 Nuclear Magnetic Resonance 187
  10.3.2 Mass Spectrometry 192
  10.3.3 Liquid Chromatography–Mass Spectrometry 194
  10.3.4 Gas Chromatography–Mass Spectrometry 196
10.4 Analysis of metabolomic data 197
  10.4.1 NMR-Based Metabolomics Data 197
  10.4.2 MS-Based Metabolomics Data 199
  10.4.3 Contextualization of Metabolomics Data 200
  10.4.4 Annotation/Identification 200
10.5 Conclusion 200
References 201

10.1 OVERVIEW

The publication of the draft human genome on February 15, 2001, [1] was a significant milestone in the understanding of human disease and diagnosis. An enormous repository of information on the basic building blocks of the human organism had been made available to researchers, allowing groundbreaking research into the relationship between genes and inherited diseases and disease susceptibility. However, by exploring the linkage between genetic influence and the overall phenotype, it was demonstrated that sequencing the genome is a starting point in the analysis of the complexity of an organism.

Genetic events propagate from DNA (the genotype) through transcription and translation to ribonucleic acids and proteins [2]; these in turn have an effect on the small molecules that make up the bulk of a cell. The combination of these levels creates the phenotype [3]. Crucially, extracellular and extraorganismal signals propagate via small molecules, proteins, and nucleic acids, effecting transcription, translation, post-translational protein modification and chemical reactions with other small molecules.

The discipline of genomics was formed to exploit the plethora of genetic information available both on human disease, including single-nucleotide polymorphisms and epigenetic modifications within the human genome, and from genomic data from other human health–related organisms such as pathogens. It has already had significant impact on health research, via genome-wide association studies [4–6] and SNP analysis [7–9].

Following on from these efforts, post-genomic technologies have developed to exploit the enormous quantity of information available at the protein and metabolite levels, known as proteomics and metabolomics, respectively.

10.1.1 The Metabolome and Metabolomics

The metabolome was first described by Oliver and colleagues in 1998 [10], during their pioneering work on yeast metabolism, as the complete complement of small molecules in a biological system or fluid. It therefore follows that the discipline of metabolomics is the analysis of the metabolome: the small molecules in a biological sample. Metabolomics is also referred to as metabonomics or metabolic profiling by some groups, but *metabolomics* is the term most commonly used and is therefore used here.

Use of gas chromatography–mass spectrometry for metabolite profiling of clinical urine samples is possibly the earliest application of what we would now describe as metabolomics, dating from the 1970s [11–13]. The application of nuclear magnetic resonance spectroscopy to profile clinical samples was pioneered by Nicholson and others in the 1980s [14–16]. Later, mass spectrometry coupled with both liquid chromatography [17] and capillary electrophoresis [18] came to be applied to metabolomics to round out the analytical techniques available.

10.1.2 Biochemistry of the Metabolome

Unlike the genome (comprising only 4 different nucleotides) and the proteome (comprising 20 different amino acids), which have a restricted set of building blocks (excluding epigenetic and post-translational modifications, which tremendously complicate these disciplines), metabolites, as individual molecules, include an enormous diversity of organic structures and functional groups that allow them to perform their particular functions. Indeed, these building blocks—amino acids, phosphosugars, nucleotides and their precursors, and degradation products—are all metabolites.

While many metabolites possess functions intrinsic to their structures, they do not exist in isolation *in vivo*. Metabolites are the substrates of enzymes and are chemically modified in anabolic and catabolic processes. Substrates may enter cells from the extracellular matrix and be transported out again as secreted molecules. They can also exist in chemical equilibrium with other molecules and convert readily between two states. For ease of reference, a series of chemical reactions mediated by enzymes in a cell are described as pathways. They include the most basic functions of any living cells, including

- Energy generation via glycolysis, the tricarboxylic acid cycle, the pentose phosphate pathway, and fatty acid oxidation, coupled with the electron transport chain; energy storage via gluconeogenesis; and fatty acid synthesis or glycogen synthesis.
- Synthesis of macromolecules via DNA replication, tRNA synthesis, and protein synthesis.

Understanding these pathways is critical to the interpretation of metabolomics data. The reader is encouraged to refer to biochemistry textbooks for help in interpreting the data [19–21].

Pathways themselves do not exist in isolation, but the inputs and outputs of each pathway are themselves the products of other pathways such that all pathways are interconnected in a complex network (see Figure 10.1) that constitutes an organism's complete metabolism [22]. This provides an enormous analytical challenge because, for example, an extremely sensitive method for the detection of organic acids may be unsuitable for the detection of amine compounds. As a consequence, the researcher must either focus her efforts on a single analytical platform

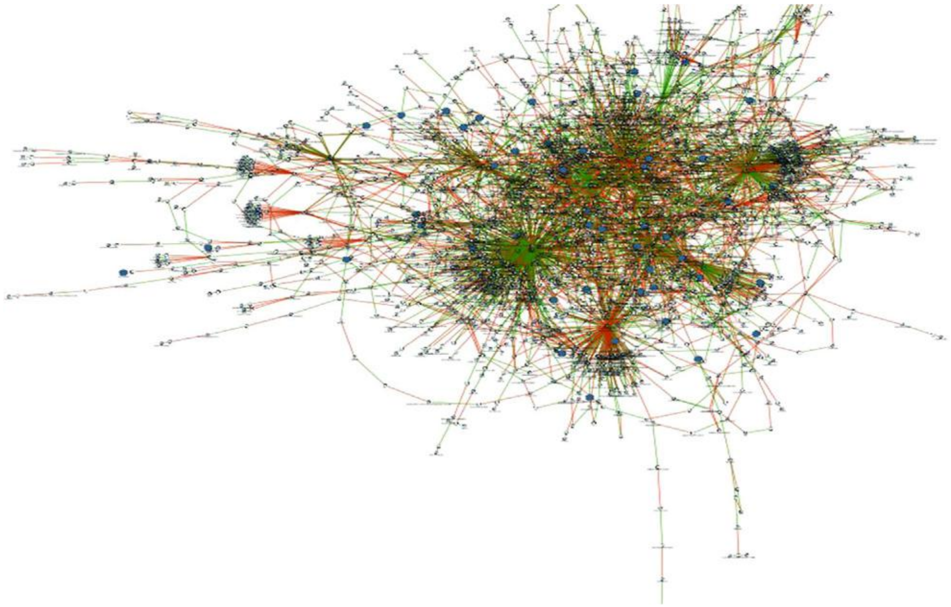


FIGURE 10.1 Complete network of plant metabolites. Metabolites are nodes; enzymatic reactions are edges. Some metabolites, notably water, are excluded from the network due to their ubiquity in reactions. Note the dense clusters surrounding some nodes, which are the consequence of highly promiscuous metabolites such as ATP. *Source: Image generated using MetExplore* [23].

that is capable of detecting a subset of the metabolome, or apply multiple methods to each sample to improve overall coverage.

The total number of metabolites in a biological system is currently the subject of considerable debate, with estimates of between a few hundred for small eukaryotes [24] to thousands in higher mammals [25], to hundreds of thousands in plants [26], with their complex array of secondary metabolites. In addition, while drugs and xenobiotics are generally not considered to be metabolites per se, analytical techniques used to detect endogenous metabolites also detect drugs because of their similar physicochemical properties; therefore, metabolomics is a powerful tool for drug mode-of-action studies.

10.1.3 Applications of Metabolomics in Clinical and Biomedical Research

While analysis of the entire human metabolome, even at the low estimate of a few thousand metabolites, is a goal that metabolomics is far away from reaching, the capability to analyze hundreds of individual small molecules or profile the characteristic signatures of the entire metabolome makes metabolomics an extremely powerful tool that can be applied to a wide variety of biological questions. In cancer science, the Warburg effect [27] of altered energy metabolism in tumor cells has recently been an important target of metabolomic research, with several groundbreaking papers on the energy metabolism of cancer [28,29]. Moderate-scale clinical metabolomics studies in the cardiovascular field, including preeclampsia [30], diabetes [31], and stroke [32], have provided clues as to common disease mechanisms. More specific studies on organs and tissues associated with significant issues in human health have been performed (e.g., liver function [33], pancreatic cancer [34], and lung function [35]).

Parasite (e.g., Barrett et al. [36], Vincent et al. [37], and Vincent et al. [38]) and bacterial (e.g., see references [39–41]) metabolomes that respond to drug action and host interactions have been extensively studied. Indeed, there is a growing interest in the microbiome and metabolomics [42,43] because, while genomics can provide descriptive analyses of the species of microorganisms and their microenvironments (e.g., the gut), the majority of metabolites are cross-species, which allows researchers to study the complete metabolome produced by a community of microbes, including its unique molecules, toxins, and other compounds.

Gas-chromatography mass spectrometry (GC-MS) is particularly powerful in the profiling of samples that contain molecules with structural similarity, such as sugars [44–48], drugs of abuse [49–51], and steroids [52–55]. While the majority of applications are targeted methods for quantitation, GC-MS has also been used for global metabolomics profiling [56–60].

Clinical samples analyzed by nuclear magnetic resonance (NMR) metabolomics include whole blood [61], reconstituted blood spots [62], plasma [63], serum [63], urine [63], saliva [64], cerebrospinal fluid [65], and digested tissue biopsies [63]. NMR-based metabolomics has also been used in lipoprotein analysis for prediction of cardiovascular risk [66,67], early identification of sepsis in critically ill patients [68,69], identification of novel inborn errors of metabolism [70], and drug toxicity [68,71].

Many organisms that model human disease have also been analyzed by metabolomics. For example *Drosophila* studies have been related to kidney function [72–74] and spleen function has been studied in mice [75]. Finally, interdisciplinary projects spanning bioengineering and orthopedic surgery have used metabolomics to optimize surfaces for the maintenance and differentiation of stem cells [76–78].

10.1.4 Experimental Design for Metabolomics

Of course, with such a diversity of applications, good experimental design is essential to achieving high-quality results. It is important to decide whether the experiment will be an unbiased study for hypothesis generation or a specific study chosen to test a hypothesis [79]. The former is the canonical biomarker discovery experiment, while the latter exemplifies validation of biomarkers in a larger cohort.

As in any experiment, a sufficiently large number of controls must be carefully chosen to eliminate bias. For health/disease studies, it is important to use positive and negative controls to separate general effects of ill health from specific disease-associated markers. Samples should also be appropriately randomized and quality control samples must be rigorously acted on if the data obtained fail to meet appropriate standards to eliminate or compensate for inter- and intrabatch variation [80]. Internal standards should also be selected to ensure effective extraction [81].

An appropriate number of biological replicates should be chosen to provide the appropriate statistical power [82]. More specifically, the larger the coefficient of variation in metabolite levels in the samples, the higher the number of replicates and the larger the minimal fold change required for statistical significance. In general, for clinical samples from small sample groups only very large variations from control levels have statistical power. Less than twofold changes in metabolite concentration in highly variable clinical datasets require very large numbers of samples (hundreds or thousands, dependent on the variation observed). While well-controlled cultured cells have low variation, a smaller number of replicates provides similar power. Technical variation in metabolomics experiments is predominantly lower than biological variation [83]. A typical experimental design for a simple pairwise comparison is shown in Figure 10.2.

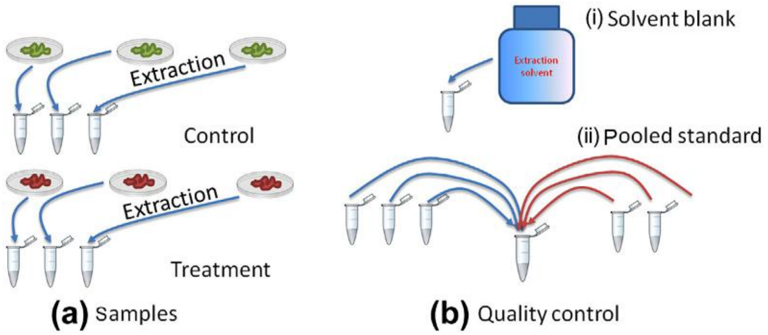


FIGURE 10.2 Simple experimental design of a pairwise metabolomics comparison. (a) Minimum three biological replicates from each state to be analyzed (replicate numbers should be carefully chosen to optimize detection of the desired fold change). An aliquot of extracted fresh medium could be added to these samples, as well as the spent medium from each biological replicate, to observe depletion of metabolites. (b) Quality control samples, including a solvent blank (i) to eliminate spurious compounds introduced in the extraction solvents and (ii) a pooled standard (i.e., a mixture of a small aliquot of the other samples) to enable tracking of compound degradation, batch effects, and instrument drift.

Once the experiment has been carefully planned, it is important to select an analytical method that will have the best chance of producing the desired outcome.

10.2 METABOLOMICS TECHNIQUES AND TECHNOLOGIES

While all "omics" technologies are capable of providing information on the state of an organism, metabolomics can also provide information on food or medium intakes as well as outputs in terms of urine, feces, spent media, and so forth. This application of metabolomics to the inputs and outputs of a system, coupled with analysis of intracellular metabolites, provides additional information. Thus it is possible to determine by comparison between spent and fresh medium that cells have depleted a compound from their environment. A commonly described example of this is the consumption of glucose from the medium. Fresh medium may contain large quantities of glucose; spent medium may not. If there is a very small pool of glucose in the intracellular metabolome, it can be inferred that the glucose has been taken up and fully utilized. Conversely, if glucose is low intracellularly and unchanged between fresh and spent medium, this may indicate that the glucose is not being used by these cells.

Metabolomics uniquely comprises a growing number of competing platforms. The primary tools for metabolomics analysis are liquid or gas chromatography hyphenated to mass spectrometry (LC-MS or GC-MS) and nuclear magnetic resonance (NMR). Capillary electrophoresis, as pioneered by Soga [18], is becoming more popular as a separation method for charged metabolites prior to mass spectral analysis. High-throughput spectroscopic profiling techniques relying on FT-IR [84,85] and Raman spectroscopy [86,87] have also become popular. Each methodology has strengths and weaknesses for particular experimental types, which should be borne in mind when designing a particular type of metabolomics experiment.

In this section, we follow the four major classifications of metabolomics experiment (fingerprinting, metabolite pool analysis, flux analysis, and metabolite profiling) that are dealt with in Birkemeyer et al.'s excellent review on metabolomic experimental paradigms [88].

10.2.1 Fingerprinting

Metabolite fingerprinting is essentially a sample stratification and classification strategy. In this type of experiment, the identification of individual compounds is not the focus, as the principle is to obtain diagnostic patterns from the data that can be used, for example, to distinguish disease status from healthy status. This strategy was pioneered in the NMR, FT-IR, and Raman spectroscopy fields, and it has been applied to several major health issues such as coronary heart disease [66] and diabetes mellitus [89].

10.2.1.1 Metabolite Fingerprinting via NMR Spectral Binning

For metabolome-wide association studies (MWAS), multivariate statistical methods are commonly used [68,90]. High-throughput NMR is performed on the samples. The data are binned and commonly, orthogonal PLS multivariate analysis is performed on the raw spectra, with correction for false discovery rate [68,90]. Once peaks that distinguish cases from controls (or similar) are identified as being important, the unknown metabolite is then identified. This is essential for biological interpretation.

Metabolic fingerprinting can be used to great practical effect—for example, in the iKnife, developed by Balog et al. [91] for real-time tissue identification during tumor surgery.

10.2.2 Metabolite Pool Analysis

The canonical targeted metabolomics experiment is essentially one where the pool of metabolite abundances is assessed (for an early example, see reference [92]). When applied to both LC-MS (e.g., see reference [93]) and GC-MS [94], absolute levels of a particular metabolite can be obtained with the addition of a 13C-, 15N-, or deuterium-labeled internal standard to the sample, accompanied by preparation of a standard curve. This obviously renders the experiment a targeted one, where only the metabolites of interest are obtained. Some efforts to produce uniformly labeled organisms have resulted in a larger number of available labeled standards [95,96], but this requires that the organism in question be culturable in bulk and have identical metabolites to those of interest, and that standard curves for abundance are still prepared for each metabolite.

Metabolite pool analyses are static—they provide no information about rates of change or movement of metabolites. They simply tell you that one or more metabolites change or do not change in abundance as a result of a stress or treatment. For drug mode-of-action studies or analysis of a knockout mutant, if the drug is an enzyme inhibitor, the metabolite prior to the enzyme is likely to be significantly upregulated, and the metabolites downstream will be downregulated or not present post-inhibition. Due to the presence of metabolic shunts that can route around biochemical pathway blockages, this is not an infallible method for analyzing the phenotype of a mutant or drug treatment, but in many cases it can produce valuable data.

Metabolite pool analyses are the standard clinical measures of metabolites. Biomarkers such as glucose for diabetes; bilirubin, used to screen for or monitor disorders of the liver; urea and creatinine used to monitor kidney function; and lactate to detect hypoxia are measured routinely as they provide an indication of dysfunction of biochemical pathways.

10.2.3 Metabolite Profiling by MS

Metabolite profiling is somewhat halfway between observing the overall phenotype by using a metabolomic measure and applying a targeted approach. Metabolite profiling focuses on identification of the largest number of compounds possible, sacrificing absolute quantitation of the sample in exchange for relative quantitation of a far greater number of compounds.

Mass spectrometers are most commonly used for this purpose, as their sensitivity to individual compounds is far greater than any other technique. It is possible to produce relative quantitation between groups of samples, such as wild type/mutant or healthy/diseased for a large number of compounds [80,97–99]. Further details on MS profiling can be found in later sections.

10.2.4 Metabolite Profiling Using NMR

NMR metabolomics can also restrict analysis to a smaller set of well-characterized metabolites. For example, the Ala-Korpela group's method measures over 200 metabolite measures: >100 directly and >100 derived. Directly measured metabolites include >20 amino acids and other clinically relevant metabolites such as glucose, creatinine, and 3-hydroxybutyrate. Also directly quantified by NMR are approximately 80 measures of lipoprotein subclass and lipoprotein content. The 99 derived measures include apollipoproteins, total branched-chain amino acids, Fischer ratio for liver fibrosis, and lipid ratios [100].

It is common in NMR metabolite profiling experiments to apply multiple pulse programs or differential extraction of samples to increase the number and type of compounds detected. The exemplar method uses three different "windows": the LIPO window for lipoproteins and the low-molecular-weight metabolites (LMWM) window, which are both performed on native serum, and the LIPID window, which is performed on extracted lipids [101].

The LIPO window is a standard 1D-^{1}H-NOESY pulse program (to be discussed later). The spectrum obtained contains highly overlapping spectra. The baseline is automatically zeroed and the peaks are aligned. The albumin background must be subtracted [103]. Regression modeling is used to quantify VLDL, LDL, IDL, HDL, and triglycerides. These models are cross-validated to NMR-independent lipid and lipoprotein analysis [101,103].

The LMWM window is a standard CPMG pulse sequence [101]. As described in the Section 10.3.2. Common 1D NMR Pulse Programs, the FIDs are acquired after a longer period, once the larger proteins and lipoproteins have stopped spinning. There is therefore much less spectral overlap. The LIPID window is obtained from extracted lipids reconstituted in deuterated chloroform containing TMS as the internal standard [101]. (See Figure 10.3.)

The results of all the different windows are brought together, and the 99 derived measures are calculated. Self-organizing maps, based on unsupervised pattern recognition of the spectra, are then used to group patients [103]. Other clinical measures are then overlaid on these maps as a visual representation of the correlation between the spectra and the variable of interest; p-values are also calculated [104].

The huge benefit of this method is that it is high-throughput, at almost 200 samples per day, which makes it ideal for large clinical cohorts. It has already been shown to be important for risk assessment for cardiovascular disease [103], prediction of diabetic complications [104], development of cognitive impairment and Alzheimer's disease [105]. Moreover, it has been used to observe the effect of birth weight, diet, physical activity, and timing of puberty on the NMR metabolites detected.

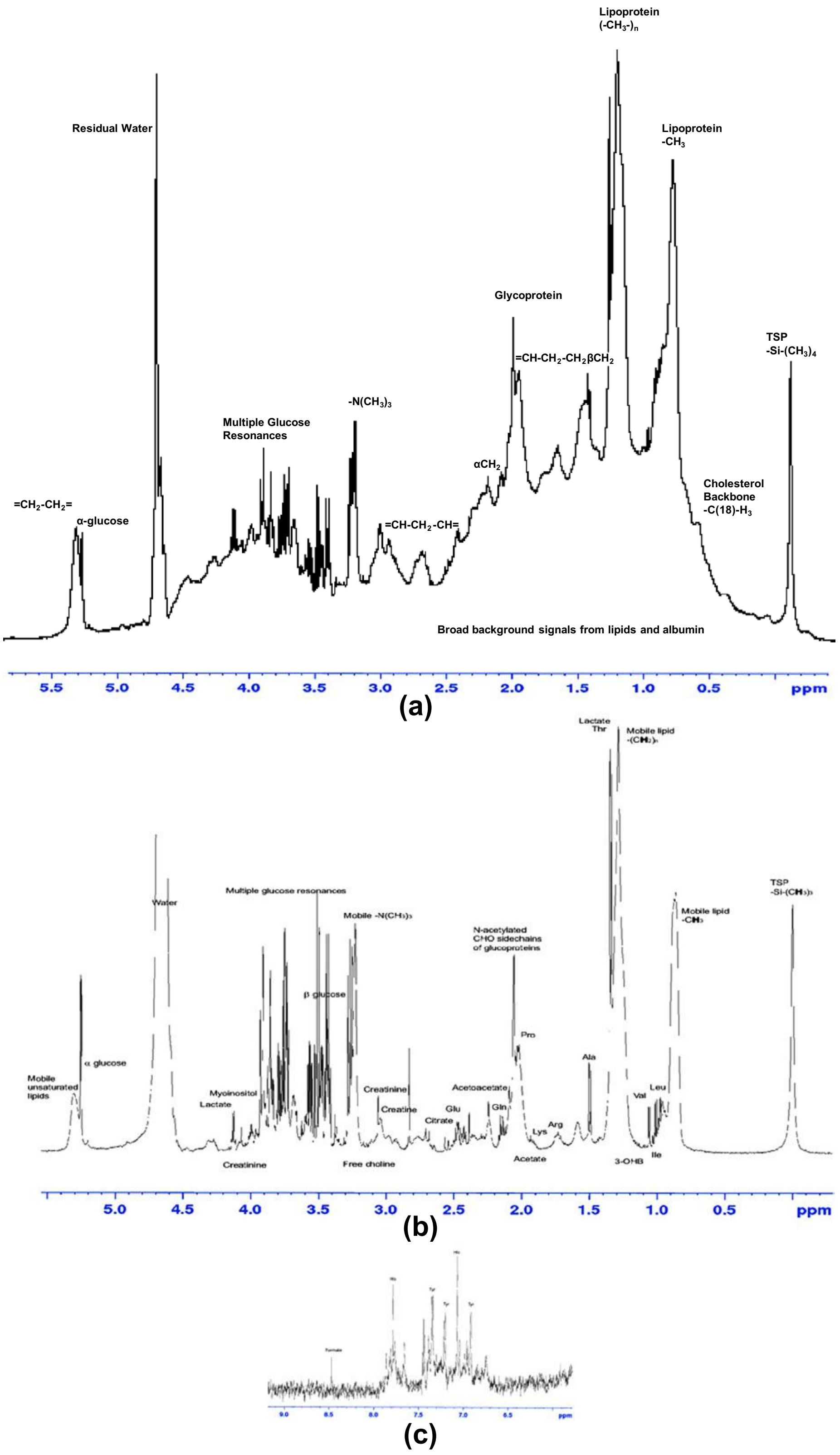


FIGURE 10.3 Example ^{1}H NMR spectra of serum obtained by the Ala-Korpela method. Samples were analyzed on a 14.1 T (600 MHz) NMR with Avance III HD console. (a) ^{1}H NOESY (LIPO window). (b) ^{1}H CPMG (LMWM window). (c) Aromatic region of LMWM window. Note that the large diffuse background in the LIPO window has been suppressed in the LMWM window. The resonances in the LIPO window can be used to determine quantitative lipoprotein profiles. The LMWM window is used to identify and quantify an array of small molecules. The inset shows the aromatic amino acid resonances. *Source: Annotation based on Soininen et al.* [102].

As the method is high-throughput and can identify and quantify multiple metabolites, it combines well with genome-wide association studies [106,107]. The more detailed the phenotyping is, the greater the ability to detect genotype-phenotype associations. This high-throughput NMR method has been used to identify genes associated with heritability of the above metabolites, especially lipids [100]. The method has also been used to identify novel genes associated with atherosclerosis [108].

10.2.5 Flux Analysis

Metabolic flux is the passage of a metabolite through a reaction system over time, and flux analysis is the combination of time-course methodologies in metabolomics and computational modeling of pathways. It is one of the most effective methods for analyzing pathways and determining the fate of a molecular substrate (see reviews in references [109,110]) and is used in fields as diverse as synthetic biology [111], antiviral therapy [112], and cancer [113].

Essentially, in stable isotope labeling an organism is provided with a food, usually medium, although some success in stable isotope labeling in higher organisms has been obtained which contains a heavy labeled version of a particular metabolite. The most common example in metabolomics is the application of ^{13}C glucose to a cell culture to trace the fate [114]. Labeled substrate is normally added in a 50% load with the unlabeled compound. Time courses investigating the rate of incorporation of a labeled substrate can be applied to inform flux balance analysis of a specific biochemical pathway with the use of appropriate software such as MzMatch.ISO [115] or MAVEN [116]. Metabolic flux analysis in the era of metabolomics can be performed using either mass spectrometry or NMR platforms, with characteristic gaps between mass peaks denoting labeled and unlabeled compounds in the case of the former.

^{13}C signals can also be detected by NMR, one of the advantages of which is that the chemical shifts are spread over approximately 200 ppm [71] (see the relevant section). Natural abundance ^{13}C (1.1%) can be difficult to detect due to the low abundance and reduced sensitivity of ^{13}C NMR. ^{13}C labeled metabolites can be made, which can help with the few metabolites that are ^{1}H NMR invisible, such as uric acid. This is a result of either a lack of protons (azide) or a very rapid exchange of protons on the metabolite with matrix protons (uric acid) [117]. It is also useful for tracking metabolite pathways.

10.3 ANALYTICAL PLATFORM

Once the study is designed and the type of experiment to be performed has been selected, it is important to determine the analytical platform. When selecting a platform for metabolomic analysis, it is important to consider its benefits and limitations prior to designing the study.

10.3.1 Nuclear Magnetic Resonance

The advantages of using nuclear magnetic resonance (NMR) for metabolomics are as follows:

- It is nondestructive and selective.
- Sample preparation is simple.
- It can be used for absolute quantitation.
- It provides information on structure
- It is independent of analyte polarity.
- It can be high-throughput and automated (associated with cost reduction).
- It can be considered an almost universal detector [118,119].
- The spectra are highly reproducible for most sample types; urine is more variable.

One-dimensional NMR is very fast, and no chromatographic preseparation or sample derivatization is required.

NMR metabolomics is also highly suitable for large-scale cohort and similar studies. Perhaps the most unique and important feature of NMR is that is can be used to identify novel compounds (generally accurate mass MS to obtain the molecular weight and 2D NMR methods are required) [120]. NMR spectroscopy is a technique that exploits the magnetic properties of atoms in order to obtain information about the structure of a molecule, and hence its identity. With its powerful quantitation capabilities [121], it is highly suited to metabolomics.

NMR detects nuclei that spin when placed in a strong magnetic field. The spin is either aligned with the magnetic field (spin-up) or opposed to it (spin-down). The nuclei in the spin-up state are in the low-energy state; those opposed to the magnetic field are in the high-energy state, as energy is required to oppose the magnetic field. Electromagnetic radiation, in the form of radiofrequency (RF) pulses, is used to push some of the nuclei in the low-energy state into the high-energy state. Once the pulse has been stopped the nuclei relax back to equilibrium. The energy is returned in the form of free induction decay (FID). These signals result from oscillation of the nuclei, which decays with time. The NMR instrument acts as both a radio transmitter and a radio receiver. The FID can be Fourier-transformed in order to convert the time domain into the frequency domain. This allows the data to be interpreted [122] (see Figure 10.3 for an example of a typical NMR spectrum of plasma using a CPMG pulse program, discussed later).

The most common nucleus observed in NMR experiments is the proton (^{1}H). Proton NMR allows absolute quantitation. It can detect almost anything with ^{1}H nuclei, including amino acids, carbohydrates, carboxylic acids, and so forth, provided it is in a high enough concentration.

There is a high abundance of hydrogen in metabolites, and NMR is very sensitive to proton signals [119]. Other commonly observed nuclei are carbons (^{13}C) and nitrogens (^{15}C) [121].

The RF must be exactly equal to the energy gap between the spin-up and spin-down states. Therefore, the higher the magnetic field strength (Tesla), the larger the difference in energy (MHz) between the two spin states (up versus down), resulting in better resolution. Sensitivity also increases with increased magnetic field strength [118]. For example, an increase from a 14.1 T/600 MHz magnet ($14.1^{3/2}=52.9$) to a 16.4 T/700 MHz magnet ($16.4^{3/2}=66.4$) results in an increase in sensitivity of 1.25 (66.4/52.9).

The first step of an NMR experiment is to design the pulse program. There are numerous predesigned sequences available with NMR software. These can be chosen from the catalog depending on the type of experiment you wish to perform and can be adapted to suit your sample type, solvent type, research question, and so forth.

A detailed guide to NMR pulses is outside the scope of this chapter. The basic pulse program begins with a delay (as this pulse will be repeated multiple times). A 90 pulse is then applied to transfer the magnetization from the z-axis to the x-y axis, where it can be observed, and to push the nuclei from the low- to the high-energy state. This is followed by the acquisition period, where the signal emitted by the precessing nuclei is recorded [121].

10.3.1.1 NMR Instrumentation

A typical NMR spectrometer setup consists of a superconducting magnet, a probe, a console, and a computer. The superconducting magnet is often a solenoid made from nobium titanium alloy. A large current flows around the loop, creating the strong continuous magnetic field required to align the nuclei. The magnet must be kept extremely cold, 4.3 K (−268.85 °C), to maintain minimal resistance (i.e., to continue acting as a superconducting magnet). This is achieved by maintaining the magnet under liquid helium in a vacuum jacket. As liquid helium is expensive and stocks are diminishing, an outer jacket containing liquid nitrogen is generally used in addition.

The sample is placed in an NMR tube and dropped into the magnet on a cushion of compressed air. The strong magnetic field causes the nuclei to become aligned or opposed to it.

The pulse program parameters are sent to the console. From here the required RF pulses are transferred to the probe. The probe is positioned in the middle of the magnetic field and contains the RF coils that go around the sample. It transmits the required RF pulses to excite the sample. It also acquires the weak FID from the nuclei in the excited sample. This is amplified and converted from an analog to a digital signal before the data are sent to the computer. Here the data can be manipulated (Fourier-transformed, phased, etc).

10.3.1.2 Experimental Techniques

Preanalytical Factors

To minimize artifacts and bias due to variation in sample collection, handling, and storage, standard operating procedures should be followed [90,123]. Care should be taken to avoid prolonged venostasis or contamination with ethanol or other substances used to clean the skin. For patients on IV drips or total parenteral nutrition, the arm with the drip in it should be avoided [19]. For toxicological metabolomics studies, consideration should be given to when the drug was administered. Fasting samples are generally recommended, as dietary intake can cause significant variation in metabolites. Time of day of sample collection should be consistent, as many metabolites, including cortisol, are subject to diurnal variation [19,123]). Where possible, multiple measures from the same patient should be used to give an average baseline.

Tube types should also be standardized. A comparison study [123] showed that the differences in NMR metabolite profile of serum compared to plasma (collected in heparin tubes) was minimal. Plasma samples collected in EDTA tubes results in huge contaminating peaks in ^{1}H NMR spectra, corresponding to free EDTA, magnesium EDTA, and calcium EDTA complexes. These peaks can potentially obscure the resonance signals of other metabolites. However, the influence of EDTA contamination was investigated and it was found that metabolites with resonances at these chemical shifts have other chemical shifts in areas not affected by EDTA and so they can still be identified and quantified [124].

For urine samples, 24 h timed urine collections are better as outputs can be normalized to time. The 24 h urine volume should be recorded and assessed for complete collection (a low urine volume may suggest an incomplete collection).

The time before centrifugation (serum-clot contact time) should be standardized. It has been found that this was the preanalytical factor, after interindividual variation, that caused the most variation in results [123]. Clotting on ice reduces this effect and is recommended [123]. There must be enough time (approximately 30 min) to allow the sample to fully clot but not so much that changes accumulate due to continued metabolism of glucose.

As much information as possible should be recorded with the sample: time of collection, time of centrifugation, temperature, time before freezing, and so forth [90,123]. Freeze–thaw cycles should be minimized by aliquoting the sample into multiple tubes. One freeze–thaw cycle has only a small affect, but multiple freeze–thaw cycles can significantly affect the metabolite profile [90,123].

Sample Preparation

For serum samples, proteins and lipoproteins can interfere with metabolite quantification; they can be removed by protein precipitation, solid-phase extraction, liquid-liquid extraction, or spectral filtering [118,125]. If not removed, they can result in variation in analytes and in analyte loss [118]. The alternative, spectral subtraction, where the sample background is subtracted, can also cause signal loss and distort the spectrum [118].

Deuterium oxide (D_2O) for locking the signal, pH buffer to minimize shifts due to pH, imidazole as a pH indicator, and azide as a biological and internal standard for referencing the chemical shift and as a quantification standard [70] are often added to each sample prior to analysis. A deuterated solvent is generally used in NMR. The magnetic field can then be locked to the ^2H in the deuterated solvent so that any changes in chemical shift of the metabolites are due to the environment of the protons within the molecule and not simply due to changes in magnetic field strength [121].

While large numbers of samples are waiting to be analyzed by NMR, the temperature should be controlled, as differences in the amount of time at room temperature can result in systematic bias [102,123]. Temperature-controlled sample changers are available for NMR.

10.3.1.3 Common 1D NMR Pulse Programs

One of the most commonly used pulses is the 1D-^{1}H-Nuclear Overhauser Effect Spectroscopy (NOESY) pulse. This is because it is a simple sequence, with few optimization requirements or hardware restrictions and good-quality water suppression is generally easily obtained [126].

The Carr-Purcell-Meiboom-Gill (CPMG) pulse program is used to identify small molecules in the presence of large proteins and lipoproteins [127]. An initial 90° pulse is applied. During the first delay, the spins begin to dephase (the nuclear spins get out of synch over only a couple of seconds). This is due to spin—spin relaxation (T2) caused by the interaction of spinning nuclei with other spinning nuclei. Their spins are different due to differences in the Larmor frequency, differences in the nuclear environment, and tiny inhomogeneities in the magnetic field. A 180° pulse is applied to rephase the spins (fast ones to the back and slow ones to the front). The FID is then acquired. This allows a longer delay between initial pulse and acquisition. Large molecules such as proteins will have stopped spinning by the time the acquisition phase is begun, so only the FIDs of the small molecules, which can spin for longer, are recorded. This pulse program, a form of spectral filtering, is very useful for serum samples, which are full of lipids and lipoproteins.

10.3.1.4 Quantitative NMR

Although NMR is described as a quantative method, errors in quantitation can be caused inadvertently if careful consideration is not given to the acquisition and processing of the samples.

Preanalytical considerations are as above (e.g., freeze—thaw). For absolute quantitation (quantitative NMR, or qNMR) a signal:noise (S:N) ratio of >100:1 is required, whereas only 10:1 is required for relative quantitiation [118]. The number of scans can be increased to increase the S:N. The increase is proportional to the square root of the number of scans added (4 scans = 2*; 16 more scans = 4*; 64 more scans = 8*); therefore, only small improvements are achieved with large numbers of extra scans [119]. A compromise must be made between analysis time and S:N achieved.

Another essential consideration for qNMR is spin-lattice relaxation time (T1). This is caused by the interaction of the nuclei with the surroundings (the lattice). It determines the time taken for the net magnetization (vector sum of the magnetic moments of the nuclei) to return to equilibrium (completely relaxed). Different protons in different metabolites will have different T1s, as they have different rates of relaxation due to differences in the nuclear environment [90]. The temperature and the solvent, because of their effects on viscosity, also affect T1. Carbon T1s are typically much longer than proton T1s. T1s can be measured by the inversion recovery method [128]. After one T1, approximately 63% of nuclei have returned to equilibrium. After five T1s, >99% have returned to equilibrium. For quantitative NMR, the pulse delay (time between pulses) must be at least five times T1 to ensure that the relative peak areas are not underestimated [118].

If this is not possible, incomplete relaxation can be corrected for, providing the T1 times have been measured in the same sample matrix [118].

10.3.1.5 Two-Dimensional NMR

Two-dimensional NMR (2D NMR) is a technique used to increase the resolution of NMR. The data are plotted against two frequency domains: the direct domain (F1), based on FT of the direct time domain (t1), and the indirect domain (F2), based on the indirect time domain (t2) obtained by increasing the delay between multiple 1D spectra. There are many methods for performing 2D NMR such as Correlaction SpectrsocopY (COSY), TOtal Correlation SpectryoscopY (TOCSY), and Nuclear Overhauser Effect SpectroscopY (NOESY). These methods are very important in metabolomics because they help determine which peaks are connected, part of the same spin system on the same molecule, when multiple metabolites are present. It is useful when there is significant spectral overlap in 1D NMR, as it increases signal dispersion. It also allows structural determination of metabolites that are too complex to work out with 1D NMR alone [63,129].

To perform a basic 2D NMR experiment, the 1D spectrum must first be optimized. Hundreds of 1D spectra are recorded, and the delay in the pulse sequence is increased with each acquisition. Resulting spectra are stacked, with the lowest t1 value at the bottom. FIDs of each acquisition are Fourier-transformed as normal to obtain the F2 (direct) frequency spectra. A new FID is formed across the stacks—in the indirect time domain (t2)—which is then Fourier-transformed to obtain the orthogonal F1 (indirect) spectrum. The indirect time domain is a result of the FIDs being sampled at discrete time intervals. F1 and F2 may be plotted against each other to observe cross-peaks [129], which are formed when the frequency of two nuclei overlap. The intensity of each cross-peak is shown on the z-axis in a topographic (contour) plot. It is determined by the efficiency of the transfer of magnetization from an excited nucleus to another, nearby nucleus. The full 2D spectrum represents all the interactions resulting from magnetization transfer: if there are a large number of paired nuclei, there will be a correspondingly large number of cross-peaks. These can then be assigned to specific positions on the metabolite [121]. Quantitation of 2D-NMR spectra is challenging [130].

A typical 2D NMR experiment consists of four stages: preparation, evolution (t1), mixing, and detection. In the preparation phase, a pulse is used to excite the nuclei. An evolution period (t1) is the delay; this is increased by a fixed amount for each pulse. During the mixing time the magnetization is transferred from the excited nuclei to other coupled nuclei within the same spin system via J-coupling or Nuclear Overhauser Effects (NOE). In the acquisition phase, the FIDs produced by the precessing nuclei are detected [129].

There are many different types of 2D NMR experiment. These mainly differ in the pulses and delays used during the mixing time to transfer the magnetization. There are two types of 2D NMR: homonuclear, where magnetization is transferred from one nucleus to another nucleus of the same type (usually ^{1}H to ^{1}H), and heteronuclear, where magnetization is transferred from one nucleus to another nucleus of a different type (usually ^{1}H to ^{13}C) [121].

Two-dimensional COrrelation SpectroscopY (COSY) [121] is a method for performing structural analysis by tracing atomic relationships through closely coupled bonds. Magnetization is transferred via J-coupling. Coupling distances can be 2-bond (geminal), 3-bond (vicinal), or (rarely) 4- or 5-bond (long range).

2D NOESY allows structural determination via relationships between atoms through space. The z-magnetization is transferred via NOE from the excited nucleus to all other nuclei in the same spin system via multiple transfers. Cross-peaks are observed if a proton is less than 5 angstroms away from another proton in the same spin system [121].

Heteronuclear single-quantum coherence (HSQC) [131] correlates proton resonances with 13C or 15N resonances that are directly coupled. This method is improved by derivatized metabolites [118]. For example, ^{13}C-glucose, can be used to perform metabolic flux analysis to identify which metabolic pathways are using the metabolite [118].

J-resolved NMR is one of the simplest 2D NMR experiments [119]. It uses a double spin echo sequence to supress strong coupling artifacts [132], so the peak splitting caused by J-coupling is reduced. This allows spectral overlap to be minimized, making the spectra easier to interpret. The lack of correlation makes J-resolved NMR less useful for structural determination; however, if the metabolites at the chemical shifts of interest are known, this is not a problem. A 2D NMR spectrum is obtained and used to identify J-coupling. The split peaks are removed and the 2D NMR is projected back as a cleaned-up version of the original 1D NMR spectrum, with only singlets [119]. The spectra are optimized and aligned, as in 1D NMR. This method has been shown to aid the identification of small molecules that would otherwise be hidden by the overlap caused by J-coupling in typical 1D NOESY or CPMG pulse programs [119]. Signals from lipoproteins and proteins are also decreased. Metabolite quantification can be more challenging in J-resolved spectra due to alterations in line shape [119,133].

A recently developed method is ultrafast 2D NMR (UF-NMR). It virtually splits the sample into multiple slices and performs a different excitation pulse on each slice. This is termed spatial encoding. Acquisition is performed on the spatial gradients within the sample. It can therefore provide 2D data in a single scan, drastically improving analysis time. Ultrafast 2D NMR has been described in detail [129]. Its benefits include increased throughput or timesaving and minimization of spectrometer instabilities that decrease the reproducibility of typical 2D NMR spectra [130]. This method has been used to look at the metabolite changes (in terms of absolute quantitation) in breast cancer cell extracts [130].

2D NMR is difficult to interpret manually, especially for complex mixtures. 2D NMR spectra may be analyzed using a variety of softwar, for example, Amix (Bruker Biospin, Germany), MetaboMiner [134], or the CCPN Metabolomics Project [135].

10.3.1.6 Other Techniques

Statistical TOtal Correlation SpectroscopY (STOCSY) [136] is used to identify correlations of spectroscopic signals with other signals. The 1D spectra of many biological replicates are compared. Peaks from the same metabolite are statistically correlated—even across multiple spin systems. Therefore, if individual peaks increase and decrease by the same magnitude as other peaks in the same sample in each replicate, it can be inferred that they are from the same metabolite, or from the molecules regulated by the same metabolic pathway [63,68].

Statistical HeterospectroscopY (SHY) [137] is used to correlate peaks from MS and NMR (or any two or more independent spectropscopic methods). Again, multiple biological replicates are analyzed by MS and NMR. If the increase or decrease of a peak is statistically correlated across the two methods, it can be inferred that they are from the same metabolite or from metabolites belonging to the same biochemical pathway [63,68].

NMR Hyphenation

Although one of the advantages of NMR is that sample preseparation is not required, separation can help in cases where a low-abundance metabolite in a sample otherwise full of high-abundance metabolites is of special interest. Generally, NMR is hyphenated to an LC system, LC-NMR. This is a technically demanding technique, as it is very difficult to shim the sample, which can result in poor spectra. Another method is simply collecting fractions from an LC using a fraction collector and running individual fractions on the NMR. Solid-phase extraction is often required before NMR spectroscopy of the fractions, as the ionic strength or pH can affect the results.

The samples can also be split so that one part is analyzed by MS and the other by NMR [63]. Depending on the MS method used, this allows determination of the exact mass, which can be used to deduce the molecular formula. Information from the retention time can also be used to putatively identify the compound. The NMR is then used to confirm the structure of the compound using the molecular formula and spectra obtained. This can be very important when cohort MS detects a metabolite that is consistently up- or downregulated in cases versus controls, but cannot be identified confidently with MS alone.

Water Suppression

The majority of samples and solvents used in metabolomics contain a large amount of water. For example, serum and urine are approximately 95% water [125]. The water signal must be suppressed or it will swamp all the other signals and can cause radiation damping [63]. There are a number of water suppression methods available that allow higher receiver gains to be used, resulting in greater sensitivity [118]. Presaturation of the water signal, the most common, is generally achieved using a long low-powered pulse at the frequency of the solvent [118]. Water suppression Enhanced by T1 (WET) and WATER suppression by GrAdient-Tailored Excitation (WATERGATE) and excitation sculpting use gradient pulses to suppress water [118,138].

Care must be taken in the quantitation of peaks close to solvent peaks, as their resonances may have been suppressed along with the solvent suppression or there may be baseline distortions resulting from the suppression [118].

10.3.1.7 *Optimization of Detection*

Acquisition time should be optimized to be short enough to prevent excess noise and long enough that the FID is not truncated [118]. The digital resolution (number of data points per Hz) should be high enough that there are at least five data points for each peak of interest [118]. Zero filling (adding zeros to the end of the FID) can be used to increase the number of data points without increasing noise [118,121]. This can allow fine coupling to be detected but can result in artifacts in the spectra. The baseline should be flat, corrected to zero, and the spectra must be properly phased in order to obtain accurate integration [118,121]. Alignment of the spectra may also be required to compensate for drift due to matrix effects [90].

10.3.1.8 *Limitations of NMR*

In comparison to mass spectrometry, NMR is less sensitive. The amount of pure material required in order to obtain a decent S:N ratio depends on the type of probe and the magnetic field strength. The probes used in metabolomics are generally cryoprobes, where the detector and preamplifier are cooled to approximately 20 K [71]. Cryoprobes have up to four times the S:N compared to room-temperature probes because the detection sensitivity of the RF coil is increased and thermal noise in the electronics is reduced [63,71,118]. Microcoil probes can also be used with small-diameter NMR tubes to increase sensitivity and are especially good for limited sample amounts [63,71,118]. The pulse is repeated multiple times on the sample in the magnet to increase the S:N ratio. For good peaks that can be used for quantitation, the sample must be carefully shimmed. A series of electromagnets are used to adjust the magnetic field in order to maximize the homogeneity of the magnetic field. This is essential for obtaining sharp peaks. If the magnetic field is poorly shimmed, the chemical shift of the nuclei will vary across the sample due to local changes in the magnetic field, resulting in broad peaks that could overlap with other peaks and be difficult to identify or quantify [121].

10.3.1.9 *Spectral Overlap*

Proton NMR is limited by the small chemical shift scale (0–10 ppm for the majority of compounds). 1D 1H NMR of mixtures can be extremely difficult. Each metabolite will have multiple signals for each proton in a different environment. Multiple signals are also obtained for stereoisomers of slowly interconverting conformations. Due the multiplicity of signals from many metabolites, NMR of mixtures can result in significant signal overlap. To overcome this problem, 2D methods can be used, but they take much longer to perform (hours compared to minutes for 1D NMR) because of the number of scans required. J-resolved NMR has the benefit of minimizing spectral overlap in a much shorter time than other 2D NMR methods [119].

10.3.1.10 *Matrix Effects*

Although NMR does not suffer from the effects of ion enhancement or suppression, as in MS, differences between sample matrices will still affect the spectra. Subtle differences in pH, ionic strength and protein content cause differences in peak position and line width. Each metabolite is differentially sensitive to these effects. In many compounds, NMR resonances of each nuclei are affected independently—making global correlation extremely difficult.

Such effects can be minimized by careful consideration of sample preparation. For example, mixing the sample with a good buffer will minimize differences in pH between samples, although for urine samples that have a wide pH range, shifts can still be seen even after buffering [63]. Note that the ionic strength of the buffer must not be so great as to adversely affect tuning and matching of the probe [63]. Also, a known concentration of imidazole can be added to act as a pH reference for the sample, as its ppm can be used to determine pH [139]. The NMR should also be regularly temperature-calibrated, for example, with deuterated methanol, to eliminate skew caused by temperature effects [140].

10.3.2 Mass Spectrometry

Mass spectrometry is an analytical technique that essentially "weighs" each molecule in a sample and plots the detected masses on a graph (Figure 10.4).

Mass spectrometers commonly used in the life sciences range from low-resolution quadrupole mass filter [141] and quadrupolar ion trap [141] instruments to high-resolution time-of-flight [142,143], orbitrap [144], and ion cyclotron resonance [145] instruments. High-resolution instruments are also those with the highest mass accuracy, but both high- and low-resolution instruments are commonly used in metabolomics applications. High resolution allows us to distinguish between, for example, tyrosine ($C_9H_{11}NO_3$, with a protonated monoisotopic mass of 182.081) and methionine sulphone ($C_5H_{11}NO_4S$, with a monoisotopic mass of 182.048) because they appear as two separate peaks (Figure 10.5). On a low-resolution quadrupole instrument, these peaks overlap and appear to be a single peak. High mass accuracy, on the other hand, provides the capability to calculate the empirical formula of a compound, based on its mass defect—that is, its deviation from unit mass resulting from the presence and number of heteroatoms in the structure, often supplemented by the intensity ratio of detected isotopes [146,147]. While this is not infallible, especially when including a wider variety of possible atoms than C, H,

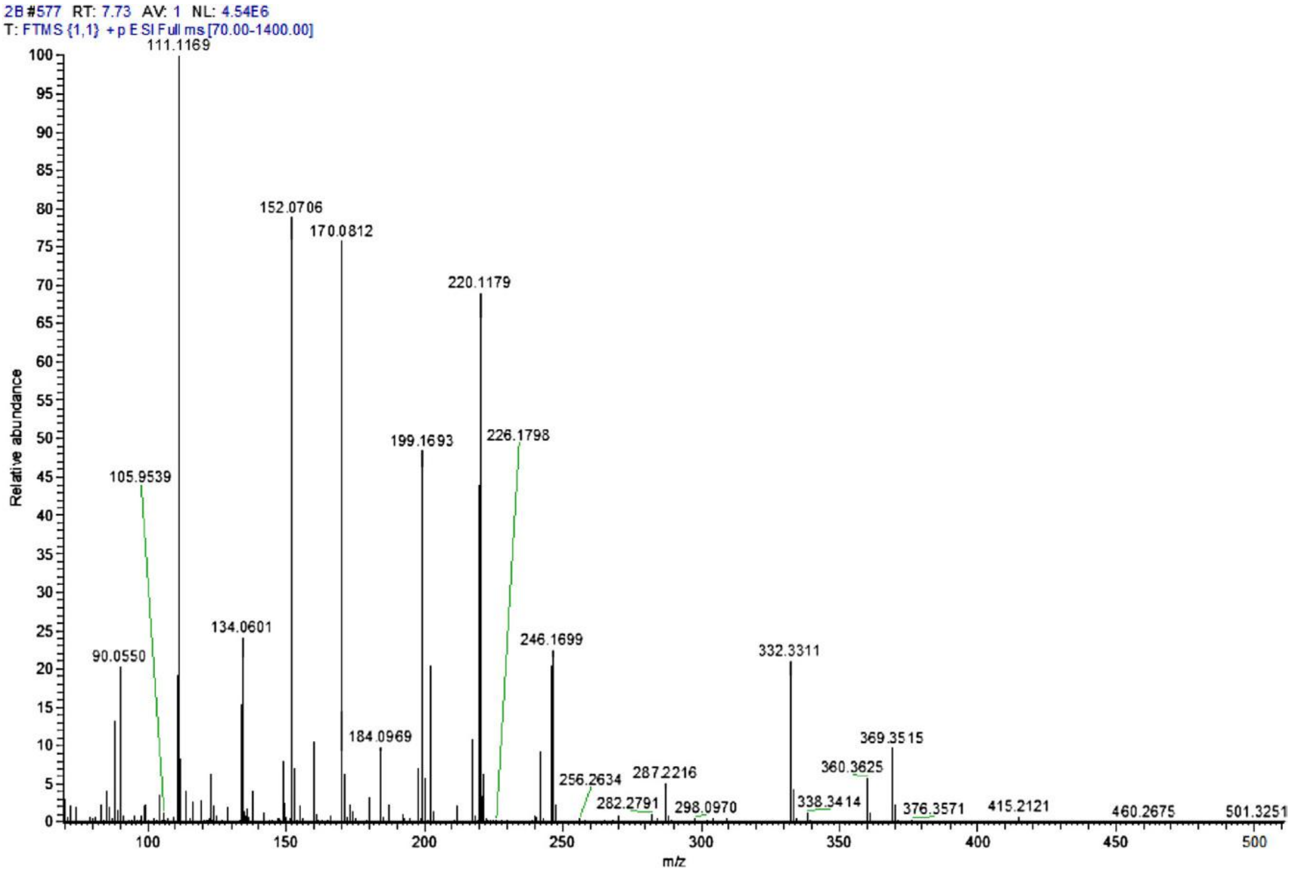


FIGURE 10.4 **Typical mass spectrum from a metabolomics experiment.** Each fine vertical peak is a detected metabolite. Mass:charge ratio is displayed along the x-axis, and relative abundance is plotted on the y-axis. Peaks are labeled with their constituent masses. The spectrum, despite being a single data point from a section of an LCMS analysis, is very complex.

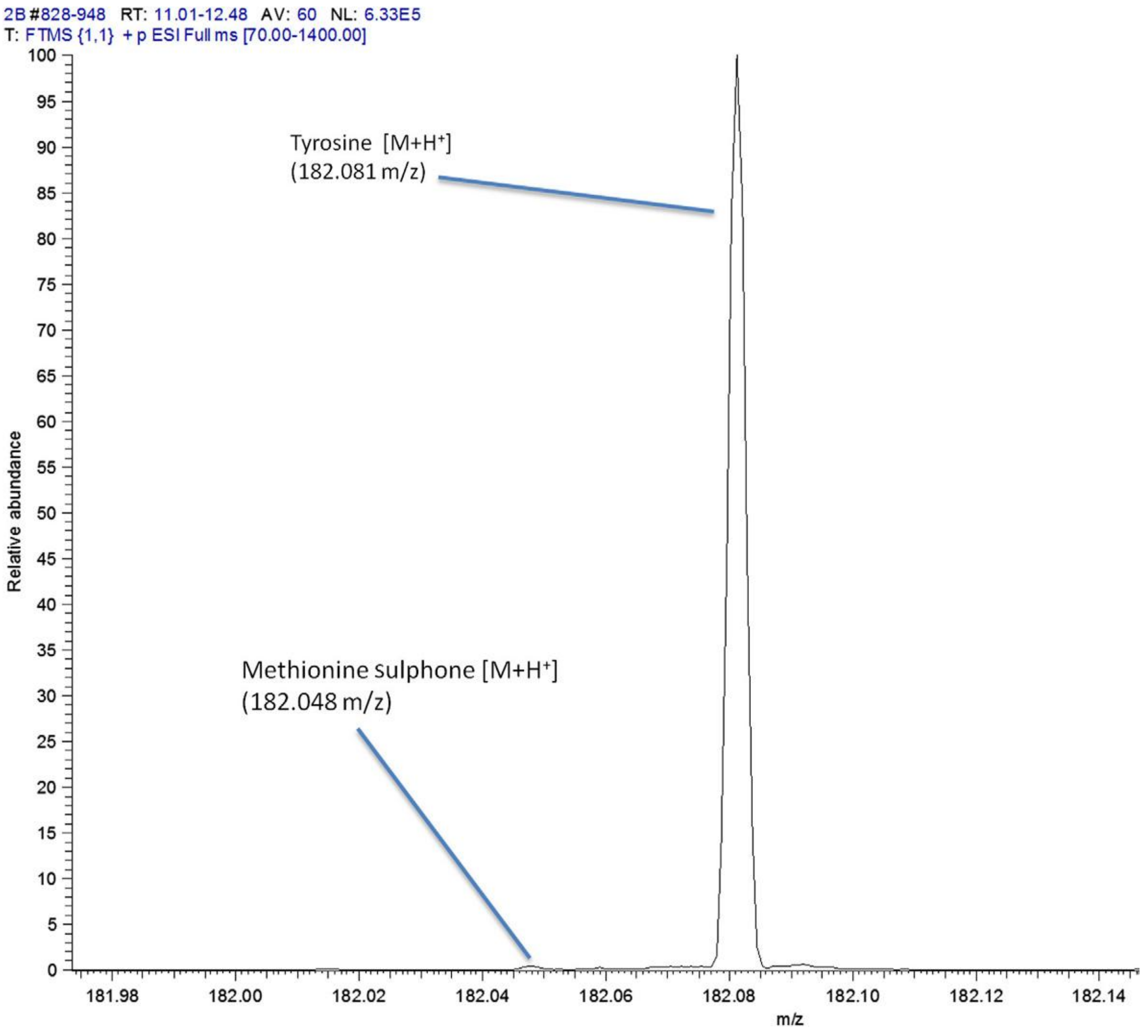


FIGURE 10.5 Compounds with similar mass. Tyrosine ($C_9H_{11}NO_3$, with a protonated monoisotopic mass of 182.081) and methionine sulphone ($C_5H_{11}NO_4S$, with a monoisotopic mass of 182.048). While methionine sulphone is very low intensity compared to tyrosine, it is still a distinct peak. A resolving power of 50,000 is amply sufficient to distinguish these compounds. Unit mass resolution quadrupole instruments would display a single peak in this instance.

O, and N [148], it is a powerful method for narrowing down the search space of possible metabolite identifications. Due to their high resolution and mass accuracy, these instruments are normally used in an untargeted manner, where an effort is made to detect, annotate, and quantify every available peak in a sample.

Unit mass resolution instruments, such as quadrupole-based mass spectrometers, often apply fragmentation to generate orthogonal information that can help them approach the diagnostic capabilities of accurate mass instruments. Of course, these fragmentation techniques are also available for high-resolution instruments for confirmation of identification. In the most common use of this type of experiment, a multiple reaction monitoring (MRM) [149,150] experiment is carried out on a triple-quadrupole instrument [151].

MRMs are based on the application of quadrupolar mass filters to exclude any ion that does not possess the appropriate unit mass. Once an ion of particular mass passes the first quadrupole, it is fragmented in a second quadrupole before the resulting fragment ions are filtered in a third quadrupole (thus, the “triple-quad” instrument type). Only ions passing through both filters strike the detector and are analyzed (see Figure 10.6). For this reason, MRM experiments are extremely sensitive and, if unique fragment ions for a particular structure can be obtained, can also be very specific. The significant drawback to this technique is that it can only be used in a targeted manner: if the experiment does not

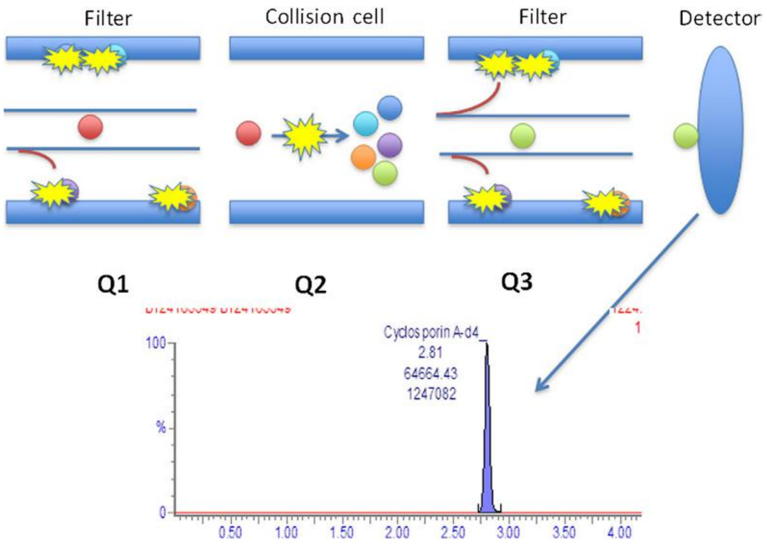


FIGURE 10.6 Multiple reaction monitoring experiment using a triple-quadrupole mass spectrometer. It was used for selective and sensitive detection of specific compounds. In this instance, the intact mass of cyclosporin is selected for quadrupole 1 (Q1), then fragmented in quadrupole 2 (Q2). The resulting ions are filtered to exclude nondiagnostic fragments in quadrupole 3 (Q3). Diagnostic fragments are allowed through the filter to strike the detector, producing a peak at the appropriate retention time in the chromatogram.

apply a particular MRM designed for a specific molecule, the molecule will not be detected.

10.3.3 Liquid Chromatography–Mass Spectrometry

Two significant drawbacks to any mass spectrometry method are isomers and ion suppression. Isomers will be discussed in more detail in Section 10.3.3.2; they are compounds with identical chemical formulas (and therefore mass) but different structures. Ion suppression (whose implications in clinical analysis are reviewed in reference [152]) occurs due to a variety of factors, predominantly modification of the solvation characteristics of the sample by an eluting compound. For this reason, it is considerably advantageous to apply a separation upstream of the mass spectrometer and usually directly coupled to it. Liquid and gas chromatography are the most commonly used methods for separating compounds prior to their introduction to the mass spectrometer. Both are based on the transient interactions of analytes with a surface (the stationary phase) while carried by a mobile phase (either a liquid or a gas). Analytes that interact strongly with the stationary phase will elute more slowly than analytes that interact weakly or not at all.

Reversed phase chromatography is the most common separation technique for LC-MS metabolomics [80,153]. It separates on the basis of hydrophobicity, so nonpolar compounds are eluted more slowly than polar compounds. Newer, but gaining more acceptance, are HILIC separations, which separate compounds on the basis of hydrophilicity and are more selective for polar compounds [72,73,154].

10.3.3.1 Sample Preparation for Liquid Chromatography–Mass Spectrometry

A significant consideration with any analytical technique is the requirement for sample preparation. Mass spectrometers are sensitive to contamination and interfering compounds, and liquid chromatography can be fouled by the use of incorrect buffers. There are, essentially, three methods for the extraction of metabolite samples: extraction of the metabolites followed by introduction of the extract to the chromatography system; extraction of the metabolites followed by lyophilization of the extract, resuspension in a suitable solvent, and introduction to the chromatography system; and cell lysis: solid-phase extraction to enrich for a particular class of compound and then processing in a manner similar to the other two methods. Exhaustive studies have been performed on optimal sample preparation methods for different sample types [155–157].

10.3.3.2 Limitations of Liquid Chromatography-Mass Spectrometry

Any analytical technique has limitations. In metabolomics, complete coverage of the metabolome is as yet not achievable. In addition to this global limitation, each platform for metabolomic analysis possesses a number of drawbacks. For critical evaluation of data, it is important to consider these drawbacks in data interpretation.

Dynamic range is a major issue in metabolomics in that the best mass spectrometers currently have a dynamic range of approximately four orders of magnitude. Given that the most abundant metabolites in a sample have millimolar concentrations (e.g., glucose has 3.5–9 mM), the minimal abundances of detected metabolites will run to the tens of micromoles while the physiological abundances of some compounds, for example, oestradiol, are in picomolar concentrations. MS sensitivity also plays a large part in dynamic range, as we will see later. It may therefore be necessary to further extract low-abundance compounds from the greater metabolome to allow detection of some classes of metabolites. Steroid hormones, for example, are generally at low levels, which makes them challenging to detect in a complex mixture.

Many compounds are labile, rapidly processed by enzymes, subject to oxidation by the atmosphere, or react chemically with other compounds in the metabolite mixture. Labile compounds such as ATP break down over relatively short periods, whereas others, like amino acids, may remain stable for months in solution. Some compounds are extremely rapidly turned over by enzymes [158]. Glycolysis is heavily optimized to rapidly process glucose to the point where at even one second post-lysis, glucose concentrations will have changed dramatically. It is therefore essential to inhibit enzymatic activity as quickly as possible without compromising the experiment, either by snap-freezing aliquots of sample or by the direct squirt method, where samples are simply pipetted, along with medium, into a cold extraction solvent. The latter obviously has the drawback that the signals from the medium will swamp the same metabolites in the cells.

While it is possible in some ways to perform extractions under a nonoxidizing atmosphere, LC-MS instrumentation is not commonly kept in an inert atmosphere and consequently the samples will still oxidize while they are waiting to be analyzed and even during separation. Glutathione and reduced glutathione are particularly prone to oxidation. Derivatization of these compounds to prevent oxidation is possible (e.g., reference [159]), but cannot be applied on a global compound basis.

Heterogeneity of metabolites was mentioned previously, but it is a significant issue for metabolomic analysis. Simply, metabolites are chemically highly diverse, so not only is there no single separation system capable of effectively retaining every metabolite; there is no mass spectrometry ionization mode capable of detecting every metabolite. Electrospray ionization [160] is soft and therefore produces molecular ions of many species, but compounds such as benzene or naphthalene will only ionize via LC-MS using atmospheric-pressure chemical ionization or photoionization [161], which is very limited for higher-molecular-weight metabolites.

Isomers are arguably the most significant problem in metabolomics. Isomers are compounds with identical empirical chemical formulas but different structures: examples are leucine and isoleucine and the hexose sugars, including glucose, fructose, and mannose. While biological samples are often known to contain only a handful of potential isomers, public chemical databases normally contain many more. These are unlikely to be present in biological samples, but they cannot normally be ruled out. There are several effective ways to distinguish between chemical isomers, but each has its limitations and there is no single effective method for distinguishing between all isomers. The most commonly applied mechanism is chromatography.

Specific chromatographic methods exist to distinguish between hexose and pentose sugars [162], sugar phosphates [163], D- and L-amino acids [164], and a plethora of other compounds. These methods are, unfortunately, specific to that group of compounds, and general methods for separating metabolites (most commonly based on reverse-phase and HILIC chromatography) are unable to separate any but the most chemically different compounds. Currently there is a growing interest in ion mobility spectrometry in the metabolomics community [165]. Ion mobility spectrometry (IMS) (see review in reference [166]) is a gas-phase method for separation of ions by charge and cross-sectional area. It is very rapid (usable in chromatographic time as an additional dimension of separation) and may be used, given a suitably high-resolution drift cell, to distinguish isomers.

The other common method for distinguishing isomers is fragmentation. Several mass spectrometry methods are available for isolating a molecular ion and breaking it into component fragments. The most common is collisionally induced dissociation (CID) or collisionally activated dissociation (CAD), which is a technique that applies thermal decomposition to a molecule. CID is usually deployed in the form of a quadrupolar collision cell or in an ion trap instrument. The former is generally a harder fragmentation than the latter, producing more small fragments. CID in an ion trap has the unique property that it can be performed in rounds, such that a precursor ion can be selected and then fragmented to produce a spectrum of fragment ions, followed by selection of a fragment that can then be fragmented, and so on. Needless to say, MS_3 (i.e., a survey scan, followed by two rounds of fragmentation) or higher provides clearer structural determination than a single round of fragmentation [120,167], but it takes longer and requires more samples. In some cases the use of MS_3 or greater is still unable to distinguish some isomers.

Sensitivity and detection limits in mass spectrometry are molecule-dependent. Certain groups, such as amines or carboxylic acid groups, are easy to ionize in a suitable buffer. Other compounds, such as uncharged molecules, aromatic species, and aldehydes, are challenging to ionize under normal conditions. Overall ionizability is determined by the totality of interactions between functional groups in a molecule. For a given amount of a compound, the signal intensity in the mass spectrometer is dependent on its chemical structure [168]. This means that the intensity of a peak for one compound is not comparable to the intensity of a peak for another compound. Relative quantitation may be performed by comparing the intensity of a compound in one sample with the same compound in another sample, assuming that matrix effects (such as ion suppression) do not compromise the signal of either peak. For absolute quantitation, it is necessary to provide a compound with identical physicochemical properties but different mass (i.e., an isotopically labeled compound) as an internal mass standard [169].

The chromatograms produced by a typical LC-MS run are highly complex. Metabolite samples, as previously described, contain hundreds to thousands of individual metabolites. This complexity is enormously enhanced by the presence of adducts and fragments [170]. Electrospray ionization, the most common method for providing a charge to a compound in biomolecular mass spectrometry, generally provides a charge by adding a proton (for positive ions) or removing one (for negative ions). Thus the proton is the simplest and most prevalent adduct in mass spectrometry. The presence of salt in the sample, however, can provide other cations to function as charge donors. Na^+, NH_4^+, and K^+ are common adducts in blood samples analyzed on systems using high-pH ammonium buffers. Electrospray ionization is relatively nondestructive in comparison to competing ionization technologies, such as electron impact ionization or atmospheric-pressure chemical ionization (APCI).

Despite this, it is very common for labile compounds to decompose during ionization. The most common losses include H_2O, PO_4^{2-} (especially from triphosphate compounds such as ATP), CO_2, and NH_3. Compounds can also dimerize in the ion source, leading to a compound double the size of the original. Additionally, it is possible for compounds to obtain multiple charges by picking up two or more protons or other cations. Spectra are further complicated by multiple adducts combining two or more of the previously described events.

An alternative methodology for metabolomics is GC-MS, which is very complementary to LC-MS and NMR platforms.

10.3.4 Gas Chromatography–Mass Spectrometry

The popularity of GC-MS is due to its high chromatographic resolution, reproducibility, high throughput, and low equipment cost. Figure 10.7 shows the separation of metabolites found in orange juice, with the most abundant peak in the chromatogram matching with the library spectrum for sucrose. The most popular method for ionization of the analytes eluting from the GC column is electron impact (EI). This is a high-energy process, resulting in the majority of molecular species fragmenting into product ions. The measured EI fragment spectrum detected by the mass spectrometer is, however, highly conserved over a wide variety of analyzers and manufacturers. The apparent lack of an identifying molecular ion is made up for by a fingerprint fragment pattern that can be matched to spectral libraries.

The most comprehensive and widely used spectral libraries are those provided by the NIST database, which contains spectra for thousands of pesticides and environmental pollutants, drugs of abuse, explosives, and clinical diagnostic molecules, among many, many others [171]. The main spectral library in the current release contains spectra from over 200,000 compounds, each having been verified by expert mass spectrometrists. This makes it possible to use GC-MS to make confident molecular identifications even in the absence of a pure standard. The Golm

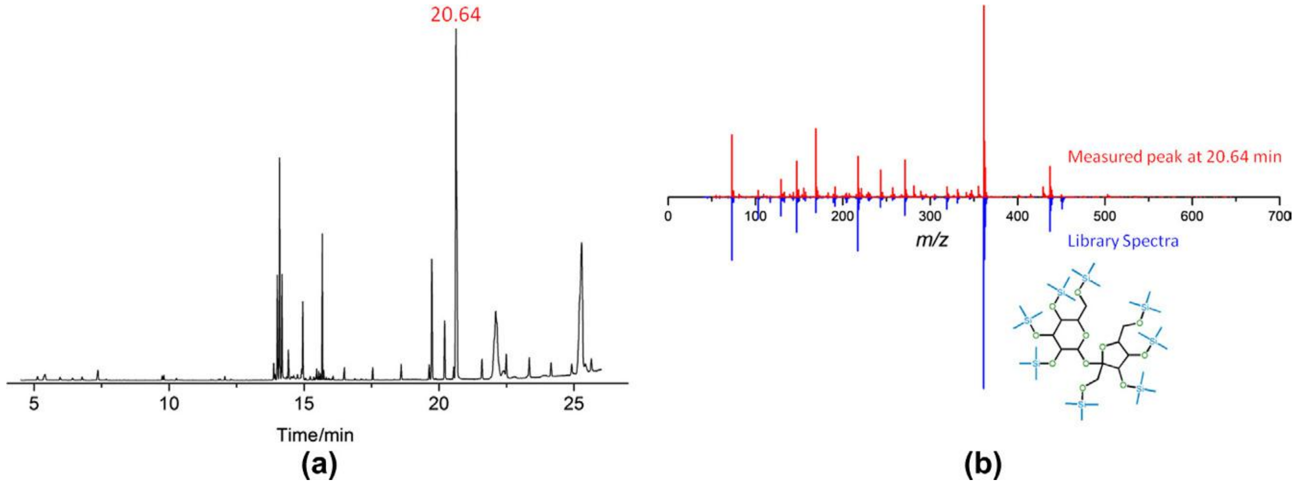


FIGURE 10.7 **Gas chromatography–mass spectrometry.** (a) GC-MS TIC for methanol extraction of orange juice derivitized with MSTFA and separated using a TraceGold TG-5MS (30 m × 0.25 mm ID, 0.25 µm film thickness; Thermo Scientific) in a Trace Ultra GC (Thermo Scientific) coupled to a ITQ 900 MS (Thermo Scientific). (b) EI mass spectrum at 20.64 min matching to the NIST library spectrum of octakis(trimethylsilyl) ether of sucrose.

Metabolome Database (GMD) is a GC-MS spectral library containing biologically active metabolites created by the scientific community [172]. Once a peak has been identified, a targeted method using single-ion monitoring (SIM) or multiple-reaction monitoring (MRM) can be used to increase sensitivity to achieve lower limits of detection (LoD) resulting in lower and more reproducible limits of quantification (LoQ). The addition of an internal standard, usually a deuterated labeled compound of interest, is used as a reference for quantitation [57,173].

10.3.4.1 Limitations of Gas Chromatography–Mass Spectrometry

To be compatible with GC-MS analysis, samples must be dissolved in a volatile solvent. The component molecules need to be volatile and thermally stable to be separated and detected. Many biomolecules contain polar functional groups such as alcohol (–OH), amide (–CONHR), amino (–NH2), carboxyl (–COOH), thiol (–SH), and enolizable ketones. Chemical derivitization results in the substitution of the acidic hydrogen, leading to a decrease in intermolecular interactions [174,175]. The derivitized compounds have an increased volatility, allowing samples that would decompose at high temperatures to be suitable for GC separation at moderate temperatures, and elute within the temperature limits of the column. The reduction in polar interactions with the column also results in better peak symmetry and resolution. Three classes of derivitization agents are silylation, acylation, and alkylation. The selection derivitization agent will depend on the compounds in the sample to be analyzed. This selection can be well known from the literature for a targeted compound, but the correct reagent may not be obvious for an unknown mixture when performing a global qualitative metabolomic analysis.

Silylation is the most commonly used derivitization reagent because these chemicals will react with both –OH (alcohols, carboxylic acids) and –NH (aminos and amines) [176], making them useful for mixtures of compounds with a biological origin. This also requires that the samples be dry in order to be free of water that would also react with the derivitization agent. Metabolite extraction and subsequent purification and derivitization can be very time-consuming. For urinary estrogens, this can involve days of sample preparation using multiple solid-phase extraction cartridges, enzymatic hydrolysis, and size exclusion chromatography as well as chemical derivitization [177]. In such cases a simpler LC-MS methodology may well be desirable.

The fragmentation pattern resulting from EI ionization can make identification of components simple if the spectrum matches a compound in the spectral library. However, this can be far from trivial if there are no high-quality matches or when compounds co-elute, resulting in a mix of fragments from both compounds. Co-elution of compounds can be overcome if there is a spectrum of a pure standard to identify which unique fragment ions are produced from the molecular ion. These unique ions can then be selectively filtered using SIM or MRM to detect only ions that confirm the identity of the eluting compound. Without a pure standard as a reference, identification of unknown compounds can be difficult, especially when there is no certainty that the molecular ion has been detected.

An experienced mass spectrometrist can interpret common fragments and use predictive tools to aid in identifying the molecule; however, this is very time-consuming and not feasible for global metabolite profiling, which can involve hundreds or thousands of unknown compounds. Chemical Ionization (CI) is a slightly less energetic process that leads to fewer, larger fragments and may allow the molecular ion to be observed, but this is not guaranteed. Some compounds, especially those with halogens or that have been derivitized with fluorine reagents, can be selectively detected using negative (−ve) CI, simplifying a complex mixture for simpler interpretation [178–180].

10.4 ANALYSIS OF METABOLOMIC DATA

10.4.1 NMR-Based Metabolomics Data

In 1D NMR, the Fourier-transformed data are represented as a spectra with chemical shift (δ) along the x-axis and intensity along the y-axis. The chemical shift is generally depicted as parts per million (ppm) rather than hertz (Hz), as this is independent of magnetic field strength.

An internal standard (IS)—tetramethylsilane (TMS), Trimethylsilylpropionic acid (TSP), 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), or their deuterated analogs— is used as a reference for 0 ppm and for determining the signal produced per proton. These molecules are chemically inert, stable, and soluble; have multiple equivalent nuclei resulting in a single sharp peak; are easily removed from the sample after measurement; and have resonances upfield of most organic compounds [181]. Ideally the IS should not evaporate at the temperatures used for analysis and should not be affected by pH or temperature.

The number of peaks present for a pure metabolite depends on the number of different environments the metabolites' nuclei are in. Because the spectra are dependent on the chemical environment of the nuclei within the molecule, NMR is a powerful method for structural determination. The chemical environment includes the shielding effect of neighboring atoms, the saturation of the molecule, and the number of neighbors [121]. The main determinant of the chemical shift is how shielded the proton is from the magnetic field. Electrons have spin; their magnetic vector opposes the overall magnetic field, shielding the nucleus. Therefore, if the electron density of a nucleus is high, the shielding effect will be higher. Electronegative atoms (such as oxygen, nitrogen, and chlorine) deshield the nuclei, resulting in a shift to the left (an increase in ppm, known as downfield). If there is more than one electronegative atom neighboring the nucleus, the deshielding effect is additive [121].

Another affect of NMR is spin-spin splitting, also known as J-coupling, which can be exploited in structure determination. If you have two protons, with different chemical environments, on adjacent carbons, their different spin states can slightly add to or slightly subtract from the overall magnetic field experienced by the other proton nucleus. If one proton is spin-up (aligned with the magnetic field), it will slightly add to the magnetic field strength experienced by the neighboring proton.

Conversely, if it is spin-down (opposed to the magnetic field), it will slightly cancel out the magnetic field strength experienced by the neighboring proton. The neighboring proton will have its peak split into two equal peaks, one at a slightly higher ppm (downfield—due to increased local magnetic field) and one at a slightly lower ppm (upfield) than the peak would be if it hadn't been subjected to these effects. The splitting pattern follows the $n+1$ rule: if a proton has one neighboring proton with a different chemical environment, it will be split in two, as just described. If it has two neighboring protons with a different environment, it will be split into a triplet with a 1:2:1 ratio. This is because 25% of the time both of these neighbors will spin up, adding to the local magnetic field; 25% of the time both protons will spin down, subtracting from the local magnetic field; and 50% of the time one will spin up and one will spin down, canceling out the effect on the local magnetic field.

Similarly, if there are three coupled protons, a quartet is observed (1:3:3:1); for four coupled protons a quintet is observed (1:4:6:4:1). The pattern follows Pascal's triangle. The area of all the split peaks added together will equal the original singlet, so quantitation is still possible. The effect is generally transmitted only through three bonds or fewer. However, coupling can be observed over more bonds if the molecule has a very rigid structure. This can therefore be used to determine the number of neighbors, helping to determine the structure [182]. The peaks are split by J Hz, where J is the coupling constant. This is independent of magnetic field, as it is caused by the magnetic field of another nucleus, not the magnet itself [121].

Chemical shift, intensity, and coupling patterns are used together to identify the metabolites in the sample. If it is suspected that a peak belongs to a certain metabolite, the other peaks known to be present for that metabolite must be observed in the spectrum, unless that part of the spectrum is overcrowded. Metabolite identification is difficult in mixtures. Note that many NMR chemical shifts are dependent on solvent, sample concentration, and other acquisition parameters (see section 10.3.3 on limitations). Metabolite identification is difficult or impossible for the wide variety of minor and trace metabolites, especially organism-specific metabolites. However, for common and abundant metabolites, automatic identification and quantification are possible using the appropriate software.

Quantitation is based on integration of the peaks and comparison with the reference peak intensity (internal standard of known concentration). The number of nuclei that contribute to the signal is also crucial [118]. For example, the ratio of a methylene (CH2) group to a methyl group (CH3) should be 2:3 for a single metabolite. Depending on whether the methyl shift or the methylene shift is used for quantitation (depending on spectral resolution), a different factor will be required.

$$\text{amount of x} = ((\text{area of x} / \text{area of IS}) \times (\text{nuclei in IS} / \text{nuclei in x})) \times \text{conc of IS}$$

The amount of x can be in grams or moles depending on the amount used for the IS. The areas are integrated using NMR software. The number of nuclei for IS depends on the structure (12 for TMS, 9 for TMS and DSS). The number of nuclei for the metabolite of interest depends on the structure and formula of the metabolite and the number of protons contributing to the signal at that specific chemical shift.

10.4.1.1 Spectral Deconvolution and Peak Picking

One method of dealing with overlapping spectra is deconvolution, which mathematically fits a Lorentzian or Gaussian peak shape to peaks suffering from partial overlap. Software that makes use of this method includes MestReNova (Mestrelab Research, Santiago de Compostela, Spain) and ACD/Labs Spectrus processor (ACD/Labs, Toronto, Cananda) [118]. Chenomx (Alberta, Canada) and Amix (Bruker Biospin, Bremen, Germany) both offer compound identification and quantification with reference to spectral libraries [70].

10.4.1.2 Binning or Bucketing

Binning, or bucketing, is a commonly used method for dealing with the crowded spectra obtained from 1D NMR of complex samples [70,118]. In its simplest form, this method involves splitting the spectrum into bins of a specific size (e.g., 0.02 ppm sections) and integrating the spectrum in the bin, regardless of whether it is from a well-defined peak or an area of overlap or is just baseline. This method cannot be used for absolute quantitation, but the integrals of the bins can be compared between samples using chemometric methods and useful information about where in the spectra differences between cases and controls (or similar) are located. Further analysis can then be performed to identify and quantify metabolites in these regions.

Care must be taken in the interpretation of binned data. As discussed previously, variations in ionic strength and pH between samples can cause the resonances of some peaks to be shifted. Slight alterations in shifts are acceptable as long as the new chemical shift stays in the same bucket; if it moves between buckets, a different bin will be increased and the original will be decreased, which can lead to misinterpretation when comparing different samples [118]. This is becoming less of a problem as alignment algorithms improve [119].

A more robust method of binning is adaptive-intelligent binning. In this case the width of the bin is not kept constant, but is adapted to take account of pH- or ionic-strength-induced shifts. It uses local minima to identify the best location to split the bins [118].

10.4.1.3 Data Normalization

Urine samples can vary significantly in the concentration of solutes depending on fluid status or the individual. Many studies use the creatinine concentration in the NMR spectra to normalize the concentrations of other metabolites [70].

The area of the full spectrum can be summed and used to normalize the solutes. However, the abundance of metabolites across different samples is not constant. For example, a diabetic patient may have significant glucosuria, which makes the other metabolites falsely low using this method [118].

By calculating the most likely dilution factor in comparison to a reference spectrum, probabilistic quotient normalization (PQN) can also be used to scale spectra to the same virtual overall concentration in order to minimize changes in NMR spectra [118,183].

10.4.2 MS-Based Metabolomics Data

When using only two-dimensional datasets, such as when using direct infusion mass spectrometry, data analysis is usually performed by binning the resulting spectrum and comparing sets of spectra using multivariate statistics [184]. These sorts of studies are generally applied in a metabolic fingerprinting process, as the lack of separation of isomers and ion suppression make this method challenging for metabolic profiling and pool analysis.

LC and GC MS produces much more complex datasets with three dimensions to consider. Each eluted metabolite will have a mass, intensity, and retention time. MRM data are easier to interpret as they are restricted to specific peaks preselected for analysis. For quantitative analysis, peaks simply have to be integrated and compared between samples or to an internal standard.

In untargeted, accurate mass LC-MS, extracting real metabolite peaks from background electrical and chemical noise is highly challenging; it requires complex chemometric and bioinformatic pipelines to sift, filter, and annotate the raw data [185]. Known metabolites can be relatively easily detected with extracted ion chromatograms (thin slices across chromatographic time, showing a single peak or a small number of peaks for a specific mass-to-charge ratio) (see Figure 10.8), but detecting all the metabolite signals

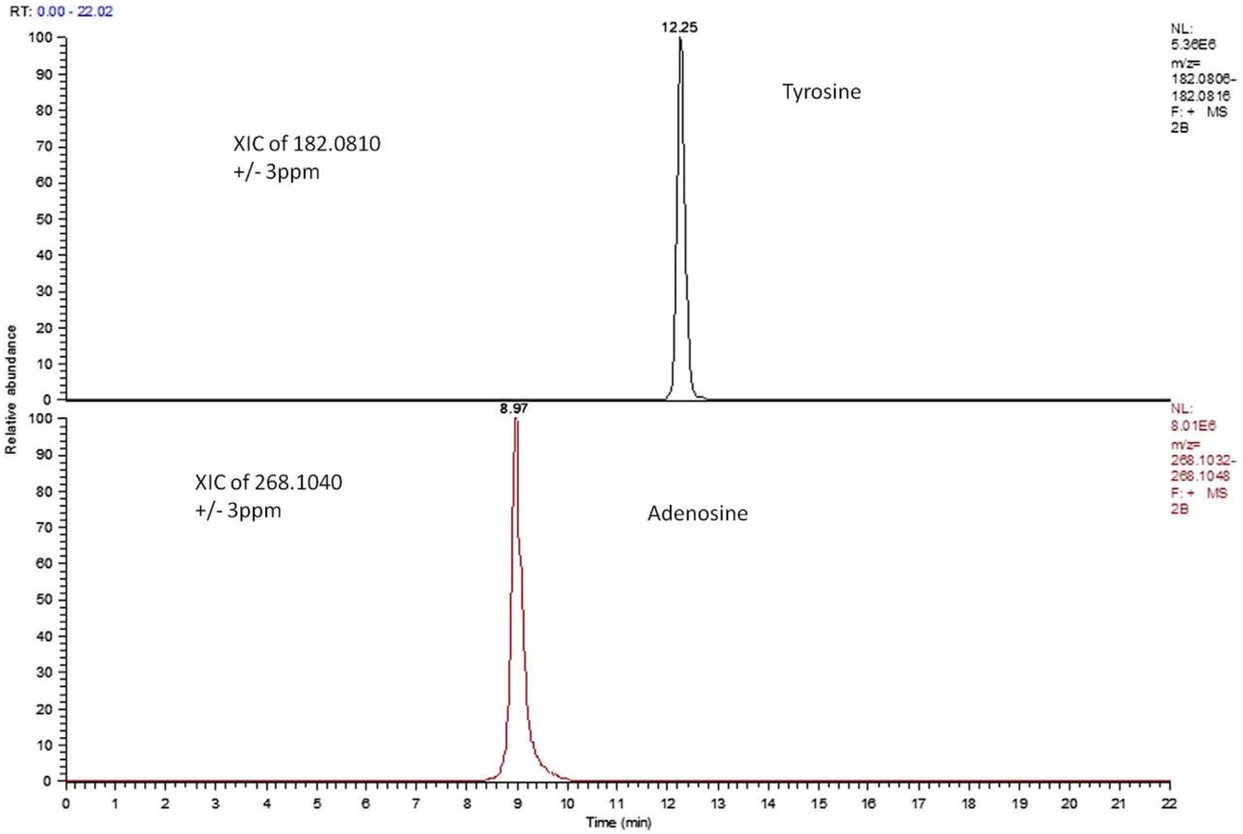


FIGURE 10.8 **Extracted ion chromatogram.** Accurate mass of tyrosine (top) and adenosine (bottom). In these instances, a 3 ppm mass tolerance is sufficient to exclude all other detected compounds from the chromatogram.

requires an unbiased method for peak picking. Several are available, including those incorporated into the XCMS [186], MzMine [187], MAVEN [116], ProGenesis CoMet, and Sieve software packages.

Once peaks are extracted, they must be processed; noisy peaks must be removed; and the entire dataset must be correlated, quantified, and identified. A good metabolomics pipeline must support all of these steps, and many pipelines are available with several alternatives for each one. It is therefore extremely important to apply the correct method for the appropriate data.

Gap filling should be performed on the data. Peak-picking algorithms are not infallible, and often peaks with nonoptimal shapes are excluded in one or more samples from a group. In cases where a peak is detected for some members of a replicate set, it is very valuable to re-search the original raw data for peaks that, rather than not being in the sample, have merely been missed by the peak detection algorithm [185]. At a later point in the processing, missing values should be annotated as n/a rather than 0, as this skews the probabilities for detected values and unrealistically represents the baseline noise level in the dataset. Missing values may also be mathematically imputed given sufficient other samples in each replicate set [188].

Normalization of metabolite data is also extremely challenging, as there are no canonical "housekeeping" metabolites available to perform normalization on. Various techniques have been applied, such as quantile normalization [189], normalization to internal standards [190], normalization to quality control samples [81], and normalization to creatinine concentration for urine samples [191]. In addition to normalization, scaling and log transformation may be applied to metabolomics datasets to reduce the dynamic range of the data and minimize skew in intensity distribution [192].

Due to the high dimensionality of the data and generally the low number of samples compared to the number of variables, it is essential that correction for multiple testing is performed on the data [193], especially if univariate statistics such as t-tests are applied to the data. The most conservative modification to the probabilities is Bonferroni correction [193]; Benjamini-Hochberg [194] correction is a less stringent method for reducing errors due to multiple testing.

10.4.3 Contextualization of Metabolomics Data

Once raw data are translated to a list of quantified metabolites, the most time-consuming and difficult part of the entire metabolomics practice starts: interpretation. Fitting the thousands of detected compounds and hundreds of significantly changing metabolites in a single experiment into a meaningful biological context is extremely challenging.

For fingerprinting analysis, patterns may be generated using multivariate statistical methods such as principal component analysis (PCA) [70], partial least squares discriminant analysis (PLS-DA) [70], or support vector machines (SVMs) [195], which will allow unsupervised (PCA) or supervised (PLS-DA and SVM) clustering of metabolomic data.

When attempting to move to the level of elucidating pathways and networks, however, it is worthwhile to begin by focusing on the most significantly changing metabolites, whether or not they are identified; after all, the most exciting discoveries are likely to be those metabolites that have not as yet been described. In general, it is helpful to examine the metabolic pathways that the compounds fit into. If most metabolites that change in one direction or another belong to the same or related pathways, this is a good indication that this pathway is important to the question being addressed.

There are several pathway-mapping packages available—for example, MetExplore [23], MassTrix [196], Pathos [197], and iPath [198]—which place already identified compounds in their constituent pathways. Pathos and MetExplore can also display up- and downregulation of the particular compounds via color coding in the generated maps. These can be very useful for observing general patterns of change in metabolites.

10.4.4 Annotation/Identification

While many of the metabolites detected may be assigned identities, it is important to classify them as either identifications or annotations depending on the surety of the assignment. According to the metabolite standards initiative [199], only compounds verified by accurate mass, MRM and retention time, or 2D NMR spectral match can be classified as identifications, while those identified by mass or substructure (fragmentation) are categorized as annotations. Although this does not cover all eventualities (especially those situations where retention time and accurate mass are not sufficient to guarantee a single identity matched to a standard), it does provide a reasonable guide to the confidence of an assignment. Of course, heuristics can be applied in context to allow greater confidence: if a number of compounds are detected in the same pathway, with the same direction of change, it is more likely that their identifications are correct.

10.5 CONCLUSION

Metabolomics is a collection of powerful tools for the analysis of phenotype, both by hypothesis generation and by hypothesis testing. Building on the strengths of the "omics technologies that came before, metabolomics uniquely comprises analytical technologies that can provide diagnostic patterns via fingerprinting, absolute quantitation of targeted metabolites via pool analysis, relative quantitation of large portions of the metabolome using metabolite profiling, and tracing of the biochemical fate of individual metabolites through a metabolic system via flux analysis. Each of these technologies is supported by the two most commonly used and powerful techniques currently available for metabolomics: mass spectrometry and NMR.

Mass spectrometry-based metabolomics techniques are the most sensitive for simultaneous analysis of a large number of compounds. While limited in quantitation capabilities without appropriate labeled standards, the sheer quantity of information available in a single LC-MS or GC-MS experiment can provide detailed information on the patterns of metabolite change in an entire metabolic network.

NMR metabolomics complements mass spectrometry. It is limited in terms of sensitivity, but is uniquely capable of elucidating molecular structure. An important additional feature of NMR is that it is quantitative, capable of providing absolute levels of detected compounds when appropriate techniques are used.

Interpreting and deriving context from complex metabolomic datasets are highly challenging and represent a major area of research. Still, great strides are being made in the integration of metabolome data with genomics and proteomics. Metabolomics offers the promise that, in the future, biochemical analysis of the entire path from genotype to phenotype will be measured and explored for new insights into biology and medicine.

REFERENCES

[1] Lander ES et al. Initial sequencing and analysis of the human genome. Nature 2001;409(6822):860–921.

[2] Crick F. Central dogma of molecular biology. Nature 1970;227(5258):561–3.

[3] Fu J et al. System-wide molecular evidence for phenotypic buffering in Arabidopsis. Nat Genet 2009;41(2):166–7.

[4] Burton PR et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661–78.

[5] Duerr RH et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 2006;314(5804):1461–3.

[6] Harold D et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 2009;41(10):1088–93.

[7] De Gobbi M et al. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. Science 2006;312(5777):1215–7.

[8] Hugot JP et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 2001;411(6837):599–603.

[9] Velaga M et al. The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (LYP) gene is a major determinant of Graves' disease. J Clin Endocrinol Metab 2004;89(11):5862–5.

[10] Oliver SG et al. Systematic functional analysis of the yeast genome. Trends Biotechnol 1998;16(9):373–8.

[11] Gates SC, Sweeley CC. Quantitative metabolic profiling based on gas chromatography. Clin Chem 1978;24(10):1663–73.

[12] Jellum E. Profiling of human body fluids in healthy and diseased states using gas chromatography and mass spectrometry, with special reference to organic acids. J Chromatogr B Biomed Sci Appl 1977;143(5):427–62.

[13] Politzer I, Dowty B, Laseter J. Use of gas chromatography and mass spectrometry to analyze underivatized volatile human or animal constituents of clinical interest. Clin Chem 1976;22(11):1775–88.

[14] Bales J et al. Metabolic profiling of body fluids by proton NMR: Self-poisoning episodes with paracetamol (acetaminophen). Magn Reson Med 1988;6(3):300–6.

[15] Gartland K et al. Pattern recognition analysis of high resolution 1H NMR spectra of urine. A nonlinear mapping approach to the classification of toxicological data. NMR Biomed 1990;3(4):166–72.

[16] Resnick L, Nicholson J, Laragh J. Calcium metabolism in essential hypertension: relationship to altered renin system activity. In: Fed proc; 1986.

[17] Dunn WB, Bailey NJ, Johnson HE. Measuring the metabolome: current analytical technologies. Analyst 2005;130(5):606–25.

[18] Soga T et al. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. J Proteome Res 2003;2(5):488–94.

[19] Burtis CA, Ashwood ER, Bruns DE, editors. Tietz Fundamentals of Clinical Chemistry. 6th ed. WB Saunders; 2008.

[20] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular biology of the cell. 5th ed. Garland, Science; 2008. 1392

[21] Stryer LBJM, Tymoczko JL. Biochemistry. 7th ed. W. H. Freeman; 2011. 1026

[22] Jeong H et al. The large-scale organization of metabolic networks. Nature 2000;407(6804):651–4.

[23] Cottret L et al. MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. Nucleic Acids Res 2010;38(Suppl. 2):W132–7.

[24] Förster J et al. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Res 2003;13(2):244–53.

[25] Wishart DS et al. HMDB: the human metabolome database. Nucleic Acids Res 2007;35(Suppl. 1):D521–6.

[26] Fiehn O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. Comp Funct Genomics 2001;2(3):155–68.

[27] Kim J-W, Dang CV. Cancer's molecular sweet tooth and the Warburg effect. Cancer Res 2006;66(18):8927–30.

[28] Chaneton B et al. Serine is a natural ligand and allosteric activator of pyruvate kinase M2. Nature 2012.

[29] Vizán P, Mazurek S, Cascante M. Robust metabolic adaptation underlying tumor progression. Metabolomics 2008;4(1):1–12.

[30] Kenny LC et al. Robust early pregnancy prediction of later preeclampsia using metabolomic biomarkers. Hypertension 2010;56(4):741–9.

[31] Lanza IR et al. Quantitative metabolomics by 1H-NMR and LC-MS/MS confirms altered metabolic pathways in diabetes. PLoS One 2010;5(5):e10538.

[32] Jung JY et al. 1H-NMR-based metabolomics study of cerebral infarction. Stroke 2011;42(5):1282–8.

[33] Lockman KA et al. Oxidative stress rather than triglyceride accumulation is a determinant of mitochondrial dysfunction in *in vitro* models of hepatic cellular steatosis. Liver Int 2012;32(7):1079–92.

[34] OuYang D et al. Metabolomic profiling of serum from human pancreatic cancer patients using 1H NMR spectroscopy and principal component analysis. Appl Biochem Biotechnol 2011;165(1):148–54.

[35] Vulimiri SV et al. Effects of mainstream cigarette smoke on the global metabolome of human lung epithelial cells. Chem Res Toxicol 2009;22(3):492–503.

[36] Barrett MP, Bakker BM, Breitling R. Metabolomic systems biology of trypanosomes. Parasitology 2010;137(9):1285.

[37] Vincent IM et al. A molecular mechanism for eflornithine resistance in African trypanosomes. PLoS Pathog 2010;6(11):e1001204.

[38] Vincent IM et al. Untargeted metabolomics reveals a lack of synergy between nifurtimox and eflornithine against *Trypanosoma brucei*. PLoS Negl Trop Dis 2012;6(5):e1618.

[39] Meyer H, Liebeke M, Lalk M. A protocol for the investigation of the intracellular *Staphylococcus aureus* metabolome. Anal Biochem 2010;401(2):250–9.

[40] Slupsky CM et al. *Streptococcus pneumoniae* and *Staphylococcus aureus* pneumonia induce distinct metabolic responses. J Proteome Res 2009;8(6):3029–36.

[41] Takahashi N, Washio J, Mayanagi G. Metabolomics of supragingival plaque and oral bacteria. J Dent Res 2010;89(12):1383–8.

[42] Calvani R et al. Gut microbiome-derived metabolites characterize a peculiar obese urinary metabotype. Int J Obes 2010;34(6):1095–8.

[43] Proctor LM. The human microbiome project in 2011 and beyond. Cell Host Microbe 2011;10(4):287–91.

[44] Mateo R et al. Capillary column gas-chromatographic identification of sugars in honey as trimethylsilyl derivatives. J Chromatogr 1987;410(2):319–28.

[45] Morrison B, Scotland CJ, Fleck A. The determination of plasma glucose in a diabetic clinic. Clin Chim Acta 1972;39(2):301–6.

[46] Vesper HW, Archibold E, Myers GL. Assessment of trueness of glucose measurement instruments with different specimen matrices. Clin Chim Acta 2005;358(1–2):68–74.

[47] Wahjudi PN et al. Measurement of glucose and fructose in clinical samples using gas chromatography/mass spectrometry. Clin Biochem 2010;43(1–2):198–207.

[48] Hannestad U, Lundblad A. Accurate and precise isotope dilution mass spectrometry method for determining glucose in whole blood. Clin Chem 1997;43(5):794–800.

[49] Hsu MC, Chen DP, Liu RH. Detection of abused drugs in urine by GC-MS. J Food Drug Anal 2009;17(4):233–45.

[50] Nair H et al. Clinical validation of a highly sensitive GC-MS platform for routine urine drug screening and real-time reporting of up to 212 drugs. J Toxicol 2013;7.

[51] Segura J, Ventura R, Jurado C. Derivatization procedures for gas chromatographic-mass spectrometric determination of xenobiotics in biological samples, with special attention to drugs of abuse and doping agents. J Chromatogr B Biomed Sci Appl 1998;713(1):61–90.

[52] Castagnetta LAM et al. Tissue content of hydroxyestrogens in relation to survival of breast cancer patients. Clin Cancer Res 2002;8(10):3146–55.

[53] Houghton E et al. Application of gas-chromatography mass-spectrometry to steroid analysis in equine sports–problems with enzyme hydrolysis. Org Mass Spectrom 1992;27(10):1061–70.

[54] Kelly GE et al. Metabolites of dietary (soya) isoflavones in human urine. Clin Chim Acta 1993;223(1–2):9–22.

[55] Liere P et al. Validation of an analytical procedure to measure trace amounts of neurosteroids in brain tissue by gas chromatography-mass spectrometry. J Chromatogr B 2000;739(2):301–12.

[56] Fiehn O, Kind T. Metabolite profiling in blood plasma. In: Weckwerth W, editor. Metabolomics. Humana Press; 2007. p. 3–17.

[57] Halket JM et al. Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. J Exp Bot 2005;56(410):219–43.

[58] Lisec J et al. Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protoc 2006;1(1):387–96.

[59] Roessner U et al. Simultaneous analysis of metabolites in potato tuber by gas chromatography–mass spectrometry. Plant J 2000;23(1):131–42.

[60] Chan ECY, Pasikanti KK, Nicholson JK. Global urinary metabolic profiling procedures using gas chromatography-mass spectrometry. Nat Protoc 2011;6(10):1483–99.

[61] Serkova NJ et al. Early detection of graft failure using the blood metabolic profile of a liver recipient. Transplantation 2007;83(4):517.

[62] Constantinou M et al. Application of nuclear magnetic resonance spectroscopy combined with principal component analysis in detecting inborn errors of metabolism using blood spots: a metabonomic approach. Anal Chim Acta 2004;511(2):303–12.

[63] Beckonert O et al. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc 2007;2(11):2692–703.

[64] Walsh MC et al. Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. Am J Clin Nutr 2006;84(3):531–9.

[65] Wishart DS et al. The human cerebrospinal fluid metabolome. J Chromatogr B 2008;871(2):164–73.

[66] Brindle JT et al. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. Nat Med 2002;8(12):1439–45.

[67] Otvos JD, Jeyarajah EJ, Cromwell WC. Measurement issues related to lipoprotein heterogeneity. Am J Cardiol 2002;90(8):22–9.

[68] Nicholson JK et al. Metabolic phenotyping in clinical and surgical environments. Nature 2012;491(7424):384–92.

[69] Serkova NJ, Standiford TJ, Stringer KA. The emerging field of quantitative blood metabolomics for biomarker discovery in critical illnesses. Am J Respir Crit Care Med 2011;184(6):647.

[70] Lutz NW, Sweedler JV, Wevers RA. Methodologies for metabolomics: experimental strategies and techniques. Cambridge University Press; 2013.

[71] Griffin JL. Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis. Curr Opin Chem Biol 2003;7(5):648–54.

[72] Kamleh M et al. Metabolomic profiling of *Drosophila* using liquid chromatography Fourier transform mass spectrometry. FEBS Lett 2008;582(19):2916–22.

[73] Kamleh MA et al. Towards a platform for the metabonomic profiling of different strains of *Drosophila melanogaster* using liquid chromatography–Fourier transform mass spectrometry. FEBS J 2009;276(22):6798–809.

[74] Malmendal A et al. Metabolomic profiling of heat stress: hardening and recovery of homeostasis in *Drosophila*. Am J Physiol Regul Integr Comp Physiol 2006;291(1):R205–12.

[75] Welthagen W et al. Comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC × GC-TOF) for high resolution metabolomics: biomarker discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice. Metabolomics 2005;1(1):65–73.

[76] McMurray RJ et al. Nanoscale surfaces for the long-term maintenance of mesenchymal stem cell phenotype and multipotency. Nat Mater 2011;10(8):637–44.

[77] McNamara LE et al. Skeletal stem cell physiology on functionally distinct titania nanotopographies. Biomaterials 2011;32(30): 7403–10.

[78] Tsimbouri PM et al. Using nanotopography and metabolomics to identify biochemical effectors of multipotency. ACS Nano 2012;6(11):10239–49.

[79] Kell DB, Oliver SG. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. BioEssays 2004;26(1):99–105.

[80] Zelena E et al. Development of a robust and repeatable UPLC− MS method for the long-term metabolomic study of human serum. Anal Chem 2009;81(4):1357–64.

[81] Kamleh MA et al. Optimizing the use of quality control samples for signal drift correction in large-scale urine metabolic profiling studies. Anal Chem 2012;84(6):2670–7.

[82] Horgan GW. Sample size and replication in 2D gel electrophoresis studies. J Proteome Res 2007;6(7):2884–7.

[83] Parsons HM et al. Spectral relative standard deviation: a practical benchmark in metabolomics. Analyst 2009;134(3):478–85.

[84] Kaderbhai NN et al. Functional genomics via metabolic footprinting: monitoring metabolite secretion by *Escherichia coli* tryptophan metabolism mutants using FT–IR and direct injection electrospray mass spectrometry. Comp Funct Genomics 2003;4(4):376–91.

[85] Keymanesh K, Darvishi MH, Sardari S. Metabolome comparison of transgenic and non-transgenic rice by statistical analysis of FTIR and NMR spectra. Rice Sci 2009;16(2):119–23.

[86] Ellis DI, Goodacre R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. Analyst 2006;131(8):875–85.

[87] Harrigan GG, Goodacre R. Metabolic profiling: its role in biomarker discovery and gene function analysis. Springer; 2003.

[88] Birkemeyer C et al. Metabolome analysis: the potential of *in vivo* labeling with stable isotopes for metabolite profiling. Trends Biotechnol 2005;23(1):28–33.

[89] Yang J et al. Discrimination of Type 2 diabetic patients from healthy controls by using metabonomics method based on their serum fatty acid profiles. J Chromatogr B 2004;813(1):53–8.

[90] Bictash M et al. Opening up the “Black Box”: metabolic phenotyping and metabolome-wide association studies in epidemiology. J Clin Epidemiol 2010;63(9):970–9.

[91] Balog J et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci Trans Med 2013;5(194):194ra93.

[92] Tweeddale H, Notley-McRobb L, Ferenci T. Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool (“metabolome”) analysis. J Bacteriol 1998;180(19):5109–16.

[93] Taylor RL, Machacek D, Singh RJ. Validation of a high-throughput liquid chromatography–tandem mass spectrometry method for urinary cortisol and cortisone. Clin Chem 2002;48(9):1511–9.

[94] Huang X, Regnier FE. Differential metabolomics using stable isotope labeling and two-dimensional gas chromatography with time-of-flight mass spectrometry. Anal Chem 2008;80(1):107–14.

[95] Bajad SU et al. Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. J Chromatogr A 2006;1125(1):76–88.

[96] Troufflard S et al. *In vivo* ^{13}C NMR determines metabolic fluxes and steady state in linseed embryos. Phytochemistry 2007;68(16):2341–50.

[97] Creek DJ et al. Metabolomic analysis of trypanosomatid protozoa. Mol Biochem Parasitol 2012;181(2):73–84.

[98] Hirayama A et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. Cancer Res 2009;69(11):4918–25.

[99] t’Kindt R et al. Metabolomics to unveil and understand phenotypic diversity between pathogen populations. PLoS Negl Trop Dis 2010;4(11):e904.

[100] Kettunen J et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet 2012;44(3):269–76.

[101] Inouye M et al. Metabonomic, transcriptomic, and genomic variation of a population cohort. Mol Syst Biol 2010;6(1).

[102] Soininen P et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst 2009;134(9):1781–5.

[103] Würtz P et al. Characterization of systemic metabolic phenotypes associated with subclinical atherosclerosis. Mol BioSyst 2011;7(2):385–93.

[104] Mäkinen V-P et al. 1H NMR metabonomics approach to the disease continuum of diabetic complications and premature death. Mol Syst Biol 2008;4(1).

[105] Tukiainen T et al. A multi-metabolite analysis of serum by ^{1}H NMR spectroscopy: early systemic signs of Alzheimer’s disease. Biochem Biophys Res Commun 2008;375(3):356–61.

[106] Ala-Korpela M, Kangas AJ, Inouye M. Genome-wide association studies and systems biology: together at last. Trends Genet 2011;27(12):493–8.

[107] Ala-Korpela M, Kangas AJ, Soininen . Quantitative high-throughput metabolomics: a new era in epidemiology and genetics. Genome Med 2012;4(4):36.

[108] Inouye M et al. Novel Loci for metabolic networks and multi-tissue expression studies reveal genes for atherosclerosis. PLoS Genet 2012;8(8):e1002907.

[109] Breitling R, Achcar F, Takano E. Modeling challenges in the synthetic biology of secondary metabolism. ACS Syn Biol 2013;2:373–8.

[110] Zamboni N, Sauer U. Novel biological insights through metabolomics and ^{13}C-flux analysis. Curr Opin Microbiol 2009;12(5):553–8.

[111] Wisselink HW et al. Metabolome, transcriptome and metabolic flux analysis of arabinose fermentation by engineered *Saccharomyces cerevisiae*. Metab Eng 2010;12(6):537–51.

[112] Munger J et al. Systems-level metabolic flux profiling identifies fatty acid synthesis as a target for antiviral therapy. Nat Biotechnol 2008;26(10):1179–86.

[113] Portais JC et al. Metabolic flux determination in C6 glioma cells using carbon-13 distribution upon [1-13C] glucose incubation. Eur J Biochem 1993;217(1):457–68.

[114] Sauer U. Metabolic networks in motion: 13C-based flux analysis. Mol Syst Biol 2006;2(1).

[115] Chokkathukalam A et al. MzMatch–ISO: an R tool for the annotation and relative quantification of isotope-labeled mass spectrometry data. Bioinformatics 2013;29(2):281–3.

[116] Clasquin MF, Melamud E, Rabinowitz JD. LC-MS data processing with MAVEN: a metabolomic analysis and visualization engine.

Curr Protc Bioinform 2012. http://dx.doi.org/10.1002/0471250953.bi1411s37 [Chapter 14:Unit 14.11].

[117] Barton RH et al. High-throughput 1H NMR-based metabolic analysis of human serum and urine for large-scale epidemiological studies: validation study. Int J Epidemiol 2008;37(Suppl. 1):i31–40.

[118] Barding Jr GA, Salditos R, Larive CK. Quantitative NMR for bioanalysis and metabolomics. Anal Bioanal Chem 2012;404(4):1165–79.

[119] Fonville JM et al. Evaluation of full-resolution J-resolved 1H NMR projections of biofluids for metabonomics information retrieval and biomarker identification. Anal Chem 2010;82(5):1811–21.

[120] van der Hooft JJ et al. Spectral trees as a robust annotation tool in LC–MS based metabolomics. Metabolomics 2012;8(4):691–703.

[121] Jacobsen N. NMR spectroscopy explained: simplified theory, applications and examples for organic chemistry and structural biology. Wiley-Blackwell; 2007.

[122] Hoult D, Ginsberg N. The quantum origins of the free induction decay signal and spin noise. J Magn Reson 2001;148(1):182–99.

[123] Teahan O et al. Impact of analytical bias in metabonomic studies of human blood serum and plasma. Anal Chem 2006;78(13):4307–18.

[124] Barton RH et al. The influence of EDTA and citrate anticoagulant addition to human plasma on information recovery from NMR-based metabolic profiling studies. Mol BioSyst 2010;6(1):215–24.

[125] Psychogios N et al. The human serum metabolome. PLoS One 2011;6(2):e16957.

[126] Mckay RT. How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: an examination of the pulse sequence components and evolution. Conc Magn Reson Part A 2011;38(5):197–220.

[127] Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. Rev Sci Instrum 1958;29(8):688–91.

[128] Becker ED et al. The choice of optimal parameters for measurement of spin-lattice relaxation times. II. Comparison of saturation recovery, inversion recovery, and fast inversion recovery experiments. J Magn Reson(1969) 1980;37(3):381–94.

[129] Tal A, Frydman L. Single-scan multidimensional magnetic resonance. Prog Nucl Magn Reson Spectrosc 2010;57(3):241–92.

[130] Le Guennec A et al. Fast determination of absolute metabolite concentrations by spatially encoded 2D NMR: application to breast cancer cell extracts. Anal Chem 2012;84(24):10831–7.

[131] Mori S et al. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. J Magn Reson Ser B 1995;108(1):94–8.

[132] Thrippleton MJ, Edden RA, Keeler J. Suppression of strong coupling artefacts in *J*-spectra. J Magn Reson 2005;174(1):97–109.

[133] Ludwig C, Viant MR. Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. Phytochem Anal 2010;21(1):22–32.

[134] Xia J et al. MetaboMiner–semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. BMC Bioinform 2008;9(1):507.

[135] Chignola F et al. The CCPN metabolomics project: a fast protocol for metabolite identification by 2D-NMR. Bioinformatics 2011;27(6):885–6.

[136] Cloarec O et al. Statistical total correlation spectroscopy (STOCSY): a new approach for individual biomarker identification from metabonomic NMR datasets. Anal Chem 2005;77:1282–9.

[137] Crockford DJ et al. Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies. Anal Chem 2006;78(2):363–71.

[138] Piotto M, Saudek V, Sklenář V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR 1992;2(6):661–5.

[139] Wishart DS. Quantitative metabolomics using NMR. TrAC Trends Anal Chem 2008;27(3):228–37.

[140] Raiford DS, Fisk CL, Becker ED. Calibration of methanol and ethylene glycol nuclear magnetic resonance thermometers. Anal Chem 1979;51(12):2050–1.

[141] Paul W, Steinwedel H. A new mass spectrometer without a magnetic field. Zeitschrift fuer Naturforschung (West Germany). Divided into Z. Nautrforsch., A, and Z. Naturforsch., B: Anorg Chem Org Chem. Biochem Biophys 1953;8:448–50.

[142] Goudsmit S. A time-of-flight mass spectrometer. Phys Rev 1948;74(5):622.

[143] Mamyrin B et al. The massreflect ron, a new non-magnetic time-of-flight mass spectrometer with high resolution. Zh Eksp Teor Fiz 1973;64:82–9.

[144] Hu Q et al. The Orbitrap: a new mass spectrometer. J Mass Spectrom 2005;40(4):430–43.

[145] Comisarow MB, Marshall AG. Fourier transform ion cyclotron resonance spectroscopy. Chem Phys Lett 1974;25(2):282–3.

[146] Fiehn O et al. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. Anal Chem 2000;72(15):3573–80.

[147] Kind T, Fiehn O. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. BMC Bioinform 2007;8(1):105.

[148] Kind T, Fiehn O. Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. BMC Bioinform 2006;7(1):234.

[149] McReynolds JH, Anbar M. Isotopic assay of nanomole amounts of nitrogen-15 labeled amino acids by collision-induced dissociation mass spectrometry. Anal Chem 1977;49(12):1832–6.

[150] Zakett D, Flynn R, Cooks R. Chlorine isotope effects in mass spectrometry by multiple reaction monitoring. J Phys Chem 1978;82(22):2359–62.

[151] Yost R, Enke C. Selected ion fragmentation with a tandem quadrupole mass spectrometer. J Am Chem Soc 1978;100(7):2274–5.

[152] Annesley TM. Ion suppression in mass spectrometry. Clin Chem 2003;49(7):1041–4.

[153] Lu W et al. Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. Anal Chem 2010;82(8):3212–21.

[154] Kamleh A, Barrett MP, Wildridge D, Burchmore RJ, Scheltema RA, Watson DG. Metabolomic profiling using Orbitrap Fourier transform mass spectrometry with hydrophilic interaction chromatography: a method with wide applicability to analysis of biomolecules. Rapid Commun Mass Spectrom 2008;22(12):1912–8.

[155] Bruce SJ et al. Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. Anal Chem 2009;81(9):3285–96.

[156] Kim HK, Verpoorte R. Sample preparation for plant metabolomics. Phytochem Anal 2010;21(1):4–13.

[157] Villas-Bôas SG et al. Global metabolite analysis of yeast: evaluation of sample preparation methods. Yeast 2005;22(14):1155–69.

[158] Koning Wd, Dam Kv. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. Anal Biochem 1992;204(1):118–23.

[159] Shimada K, Mitamura K. Derivatization of thiol-containing compounds. J Chromatogr B Biomed Sci Appl 1994;659(1):227–41.

[160] Dole M et al. Molecular beams of macroions. J Chem Phys 1968;49:2240.

[161] Robb DB, Covey TR, Bruins AP. Atmospheric pressure photoionization: an ionization method for liquid chromatography-mass spectrometry. Anal Chem 2000;72(15):3653–9.

[162] Guignard C et al. Analysis of carbohydrates in plants by high-performance anion-exchange chromatography coupled with electrospray mass spectrometry. J Chromatogr A 2005;1085(1):137–42.

[163] Antonio C et al. Quantification of sugars and sugar phosphates in *Arabidopsis thaliana* tissues using porous graphitic carbon liquid chromatography-electrospray ionization mass spectrometry. J Chromatogr A 2007;1172(2):170–8.

[164] Desai MJ, Armstrong DW. Analysis of native amino acid and peptide enantiomers by high-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry. J Mass Spectrom 2004;39(2):177–87.

[165] Dwivedi P et al. Metabolic profiling by ion mobility mass spectrometry (IMMS). Metabolomics 2008;4(1):63–80.

[166] Kanu AB et al. Ion mobility–mass spectrometry. J Mass Spectrom 2008;43(1):1–22.

[167] Rojas-Cherto M et al. Metabolite identification using automated comparison of high-resolution multistage mass spectral trees. Anal Chem 2012;84(13):5524–34.

[168] Gross JH. Mass spectrometry: a textbook. 2nd ed. Springer; 2011. p. 774

[169] Honour JW. Development and validation of a quantitative assay based on tandem mass spectrometry. Ann Clin Biochem 2011;48(2):97–111.

[170] Huang N et al. Automation of a Fourier transform ion cyclotron resonance mass spectrometer for acquisition, analysis, and e-mailing of high-resolution exact-mass electrospray ionization mass spectral data. J Am Soc Mass Spectrom 1999;10(11):1166–73.

[171] NIST standard reference database 1A2011: national institute of standards and technology.

[172] Kopka J et al. GMD@CSB.DB: the Golm Metabolome Database. Bioinformatics 2005;21(8):1635–8.

[173] Kvitvang HFN et al. Highly sensitive GC/MS/MS method for quantitation of amino and nonamino organic acids. Anal Chem 2011;83(7):2705–11.

[174] Knapp DR. Handbook of analytical derivatization reactions. Wiley-Intercience; 1979.

[175] Evershed RP, Blau K, Halket JM. Handbook of derivatives for chromatography. 2nd ed. Wiley; 1993.

[176] Pierce AE. Silylation of organic compounds. Pierce Chemical Company; 1968.

[177] Xu X et al. A liquid chromatography-mass spectrometry method for the quantitative analysis of urinary endogenous estrogen metabolites. Nat Protoc 2007;2(6):1350–5.

[178] Fine DD et al. Quantitation of estrogens in ground water and swine lagoon samples using solid-phase extraction, pentafluorobenzyl/trimethylsilyl derivatizations and gas chromatography–negative ion chemical ionization tandem mass spectrometry. J Chromatogr A 2003;1017(1–2):167–85.

[179] Pawlosky RJ, Sprecher HW, Salem N. High sensitivity negative ion GC-MS method for detection of desaturated and chain-elongated products of deuterated linoleic and linolenic acids. J Lipid Res 1992;33(11):1711–7.

[180] Peters FT, Kraemer T, Maurer HH. Drug testing in blood: validated negative-ion chemical ionization gas chromatographic–mass spectrometric assay for determination of amphetamine and methamphetamine enantiomers and its application to toxicology cases. Clin Chem 2002;48(9):1472–85.

[181] Wishart DS et al. 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J Biomol NMR 1995;6(2):135–40.

[182] Hahn E, Maxwell D. Spin echo measurements of nuclear spin coupling in molecules. Phys Rev 1952;88(5):1070.

[183] Dieterle F et al. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. Anal Chem 2006;78(13):4281–90.

[184] Allen J et al. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. Nat Biotechnol 2003;21(6):692–6.

[185] Castillo S et al. Algorithms and tools for the preprocessing of LC–MS metabolomics data. Chemometrics Intell Lab Syst 2011;108(1):23–32.

[186] Smith CA et al. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 2006;78(3):779–87.

[187] Pluskal T et al. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinform 2010;11(1):395.

[188] Hrydziuszko O, Viant MR. Missing values in mass spectrometry based metabolomics: an undervalued step in the data processing pipeline. Metabolomics 2012;8(1):161–74.

[189] Sung Min P, Hae Won L, Young-Ran Y. Quantile normalization approach for liquid chromatography–mass spectrometry-based metabolomic data from healthy human volunteers. Anal Sci 2012;28:801.

[190] Nordström A et al. Nonlinear data alignment for UPLC-MS and HPLC-MS based metabolomics: quantitative analysis of endogenous and exogenous metabolites in human serum. Anal Chem 2006;78(10):3289–95.

[191] Chen Y, Shen G, Zhang R, He J, Zhang Y, Xu J, et al. Combination of injection volumes calibration by creatinine and MS signals' normalization to overcome urine variability in LC-MS-based metabolomics studies. Anal Chem 2013;85(16):7659–65.

[192] van den Berg RA et al. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC genomics 2006;7(1):142.

[193] Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics 2006;2(4):171–96.

[194] Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990;9(7):811–8.

[195] Mahadevan S et al. Analysis of metabolomic data using support vector machines. Anal Chem 2008;80(19):7562–70.

[196] Suhre K, Schmitt-Kopplin . MassTRIX: mass translator into pathways. Nucleic Acids Res 2008;36(suppl 2):W481–4.

[197] Leader DP et al. Pathos: a web facility that uses metabolic maps to display experimental changes in metabolites identified by mass spectrometry. Rapid Commun Mass Spectrom 2011;25(22):3422–6.

[198] Letunic I et al. IPath: interactive exploration of biochemical pathways and networks. Trends Biochem Sci 2008;33(3):101–3.

[199] Sumner LW et al. Proposed minimum reporting standards for chemical analysis. Metabolomics 2007;3(3):211–21.

This page is intentionally left blank

Chapter 11

Stem Cells

Komal Loya
College of Medical Veterinary and Life Sciences, Institute of Cardiovascular and Medical Sciences,
BHF Glasgow Cardiovascular Research Centre, University of Glasgow, UK

Chapter Outline

11.1 Overview 207
11.2 Stem Cell Types 208
  11.2.1 Adult Stem Cells 208
  11.2.2 Fetal Stem Cells 208
  11.2.3 Parthenogenetic Embryonic Stem Cells 209
  11.2.4 Germline-Derived Pluripotent Stem Cells 210
  11.2.5 Induced Pluripotent Stem Cells 210
11.3 Stem Cell Culture Methods 211
11.4 Molecular Signatures of Stem Cells 212
  11.4.1 Epigenetic Modifications of Stem Cells 212
  11.4.2 Chromatin Modifications of Stem Cells 214
  11.4.3 Alteration of Gene Expression in Stem Cells 215
  11.4.4 Genomic Imprinting in Stem Cells 215
11.5 Epigenomics of Stem Cells 215
  11.5.1 Epigenetic Variation among Pluripotent Stem Cells 215
  11.5.2 Importance of Epigenomics in Pluripotent Stem Cells for Regenerative Medicine 216
11.6 Pharmacogenomics of Stem Cells 218
  11.6.1 Generation of Ethnically Diverse Pluripotent Stem Cells for Pharmacogenomic Studies 218
  11.6.2 Creating Stem Cell Banks for Pharmacogenomic Studies 219
  11.6.3 Use of Stem Cells for Toxicogenomic Studies 219
11.7 Potential of Stem Cells 219
  11.7.1 Use of Stem Cells for Disease Modeling 219
  11.7.2 Use of Stem Cells for Toxicity Modeling 221
  11.7.3 Use of Stem Cells for Cancer Therapy 226
11.8 Future Prospects 226
Abbreviations 227
References 228

11.1 OVERVIEW

Regenerative medicine is a fast developing field of research that aims to contribute new therapies to the treatment of various diseases. Stem cells play a major role in regenerative medicine [1] and human cells are usually the key focus. Human cells come in a variety of types, such as adult stem cells, somatic cells, and the more conventional embryonic stem (ES) cells. These cells can potentially be used to regenerate human cells, organs, and tissues and contribute toward restoration of normal function after disease or injury [2]. The Russian histologist Alexander A. Maximov first proposed the term “stem cell” in 1908, at the annual Congress of the Hematologic Society in Berlin. He coined this term based on the process by which the stem of a tree gives rise to various branches. However, it was not before the 1960s, that scientists realized the potential of these cells.

Stem cells are usually found in multicellular organisms and have the unique capacity for self-renewal. They undergo mitotic cell division and can differentiate into different cell types. During embryonic development, stem cells can form cells of all three germ layers such as mesoderm, endoderm, and the ectoderm. In adults, together with progenitor cells, they can contribute as a repair system for the body as well as maintain the normal turnover of regenerative organs such as blood and skin. This characteristic of stem cells makes them an important tool for potential use in medical therapies [3]. Based on their potency or differentiation ability, stem cells can be divided into five main categories (Table 11.1).

Totipotent Stem Cells. These are stem cells that have a capacity to differentiate into all cell types in an organism, including extra-embryonic tissues. They are usually formed during sexual and asexual reproduction and have the potential to develop into a fully functional embryo and then a fully functional organism. The zygote, which is formed when an oocyte is fertilized by a sperm, falls into this category.

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00011-6
© 2014 Elsevier Inc. All rights reserved.

TABLE 11.1 Characterization of Stem Cells Based on Their Potency and Differentiation Potential

Type	Characteristic	Example
Totipotent	can differentiate into all cell types including extra-embryonic tissue, ability to from a fully functional complete organism	cells of the zygote
Pluripotent	can differentiate into all three germ layers, ability to from any of the fetal or adult cell types	ES cells, iPs cells, gPS cells, pES cells, and fetal stem cells
Multipotent	usually consist of progenitor cells, ability to differentiate into only a limited number of cell types	adult stem cells such as HSCs and neural stem cells
Oligopotent	usually consist of cells that reside in the tissue, ability to terminally differentiate into cells of a specific tissue	stem cells present on the mammalian ocular surface
Unipotent	ability to differentiate into a single type of cells	progenitor cells present during postnatal prostate development

Pluripotent Stem Cells. These cells can differentiate into all the three germ layers such as mesoderm, endoderm, and ectoderm. They can differentiate into any of the fetal or adult cell types, but they do not have the ability to form embryos (e.g., embryonic stem cells).

Multipotent Stem Cells. These are usually progenitor cells that are capable of differentiating into only a limited number of cell types. Somatic cells such as neural, bone marrow–derived, or hematopoietic stem cells (HSCs) fall into this category [4].

Oligopotent Stem Cells. These stem cells usually reside in the tissue and can terminally differentiate into cells of a specific tissue [5]. For example, a recent study reported the widespread presence of oligopotent stem cells on the mammalian ocular surface [6].

Unipotent Stem Cells. Unipotent stem cells usually have the ability to differentiate into a single cell type, despite having the property of self-renewal. A recent study successfully highlighted the contribution of unipotent progenitor cells during postnatal prostate development [7].

11.2 STEM CELL TYPES

Based on the previous classification, our body harbors different stem cell types, which are discussed in more detail in the following subsections.

11.2.1 Adult Stem Cells

Adult stem cells are present in fully developed tissues and play a role in maintenance and growth of these tissues throughout their life. They are also termed tissue progenitor cells and are usually multipotent. Adult stem cell or progenitor cell populations are present in many adult tissues such as bone marrow/peripheral blood (hematopoietic stem cells), gastrointestinal tract, brain, skin, muscle, nerve cells, liver, eye, pancreas, and dental pulp [8,9]. They are usually tissue-specific and can only differentiate into cell types associated with the tissue they reside in. They are controlled by their microenvironment or niche [8,10].

Of all the various cell types, hematopoietic stem cells, which are CD34+, are extensively studied and used in therapeutic applications. Exceptions to the tissue-specific characteristic of adult stem cells are the multipotent mesenchymal stem cells (MSCs), which are derived from the bone marrow stroma. They are referred to as multipotent adult progenitor cells and are able to differentiate into a variety of tissues such as neuronal, adipose, muscle, gut, liver, lungs, and spleen, but not bone marrow or gonad cells. They are also able to differentiate developmentally when injected into a blastocyst [8,11]. Adult stem cells are unique in their differentiation ability, such that one type of adult stem cell can differentiate into another type of tissue under suitable growth conditions. This property has been termed as “plasticity” [9].

11.2.2 Fetal Stem Cells

Fetal stem cells are present in the fetus, are pluripotent, and contribute in the initial development of all tissues before birth [9]. They are derived either from fetal tissue or from supportive extra-embryonic tissue; these include umbilical cord blood (UCB), amniotic fluid (AF), Wharton’s jelly, amniotic membrane, and the placenta. As these tissues are normally discarded after birth, their stem cell populations can be used without any ethical concerns. They are similar to embryonic stem cells and multipotent adult stem cells in their capacity for self-renewal and differentiation. Recent studies have come up with various strategies to exploit the therapeutic potential of these cells [12]. Each fetal stem cell type is discussed in more detail next.

11.2.2.1 Umbilical Cord Blood Stem Cells

Umbilical cord blood (UCB) multipotent cells are present in umbilical cord blood. They are genetically similar to the newborn child [9]. Almost 1% of the total population of UCB cells is CD34−, CD34+ (which is a marker for hematopoietic stem cells) and CD38− (which is a marker for immune cells such as white blood cells). These cells have an established capacity for differentiation and self-renewal. Moreover, cord blood is a rich source of MSCs, which support the *in vivo* expansion of HSCs. Cord blood cells, are the second main source of MSCs after bone marrow, and they harbor a subpopulation of cells that express neuronal markers and may act as progenitors of neuronal cells [12,13].

11.2.2.2 Amniotic Fluid Stem Cells

Amniotic fluid (AF) harbors a heterogeneous population of cells from all three germ layers consisting of several partially differentiated cell types and progenitors. These cells are mostly of epithelial origin and are derived from the fetus or the inner surface of the amniotic membrane. AF cells isolated from earlier gestation periods express higher levels of mesodermal and endodermal markers compared to their later counterparts, but the expression of ectodermal markers remains the same. They also contain MSCs and other multipotent stem cells that express stem cell-specific markers such as Oct4 (octamer-binding protein 4), SSEA4 (stage-specific embryonic antigen 4), and Nanog. AF-MSCs are able to differentiate into a variety of cells such as adipocytes, chondrocytes, osteocytes, hepatocytes, neural cells, and cardiomyocytes [12].

11.2.2.3 Wharton's Jelly Stem Cells

Wharton's jelly cells are derived from the outer region of the human umbilical cord, which is known as Wharton's jelly and can be classified into three types: umbilical cord matrix stem cells (UCMSCs), umbilical cord perivascular cells (UCPVs), and umbilical cord stromal cells (UCSCs). UCMSCs are known to facilitate muscle generation; UCPVs lead to a population of MSCs that differentiate into bone nodules, and UCSCs differentiate into several different lineages except neuronal type 1 cells [12].

11.2.2.4 Amniotic Membrane Stem Cells

The amniotic membrane is derived from the epiblast and consists of three layers: epithelial cells (inner layer), mesenchymal cells (outer layer), and an intermediate layer that lacks any cellular component. Since these cells originate from the epiblast, they display several stem cell characteristics and can differentiate into cardiomyocytes, hepatocytes, and endothelial cells *in vitro* [12].

11.2.2.5 Embryonic Stem Cells

Embryonic stem (ES) cells are pluripotent and are derived from the inner cell mass of a blastocyst. That is, they have a capacity for indefinite self-renewal apart from being able to maintain themselves in an undifferentiated or pluripotent state. They also have the unique ability to colonize cells of different lineages, which includes the germline, when they are reintroduced into the blastocyst. Apart from their *in vivo* pluripotent potential, these cells can maintain their ability to self-renew and differentiate *in vitro* [14−16]. ES cells also express the enzyme alkaline phosphatase. They are commonly used and derived from the inner cell mass (ICM) of the blastocyst (Figure 11.1). During embryonic development, ES cells are formed from the cleavage of the zygote, which usually occurs before the formation of the blastocyst. They are usually considered to be totipotent and can therefore differentiate into any kind of cells in the body and can also form a complete organism [9].

ES cells have been derived from a variety of mammalian species over recent years. The first ES cells in mammals were derived from the blastocyst of a mouse in 1981. They had the ability to differentiate into cells of all the three germ layers, were able to form teratomas when injected into immunocompromised mice, and were also karyotypically normal and had the ability to form chimeras [17,18]. An ES cell line has also been successfully established from the Rhesus monkey, which is a nonhuman primate. The cells were shown to satisfy all the characteristics that define pluripotency, such as self-renewal, ability to differentiate into all three germ layers, maintenance of a normal karyotype, and ability to form tumors when injected into immunocompromised mice, apart from showing the presence of the enzyme alkaline phophatase [19]. In the case of humans, the first ES cell line was successfully derived in 1998 from frozen embryos that were obtained from *in vitro* fertilization. They satisfied all the properties of pluripotency and have greatly contributed to stem cell research in recent years [20]. Blastocysts of other species such as rats have also been used to derive ES cells, which were successfully able to contribute to germline transmission and chimera formation [21,22].

11.2.3 Parthenogenetic Embryonic Stem Cells

Parthenogenetic embryonic stem (pES) cells were first described in mice by parthenogenetic activation of haploid embryos using chemicals such as ethanol [23]. Parthenogenesis is a process usually found in lower organisms to generate a live organism directly from an oocyte without activation by a sperm. The mouse embryos

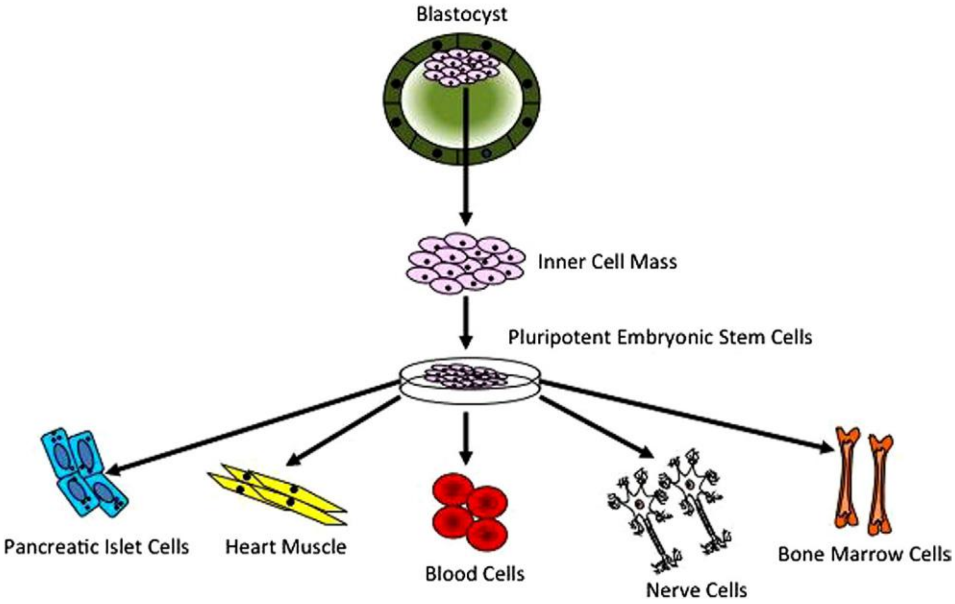


FIGURE 11.1 **Embryonic Stem (ES) Cells.** ES cells are derived from the inner cell mass of the blastocyst. When cultured under suitable conditions, they can be differentiated into cells of the three germ layers: mesoderm, endoderm, and ectoderm. Heart muscle, bone marrow, and blood cells represent the mesoderm; the pancreatic islet cells represent the endoderm; and the nerve cells represent the ectoderm.

generated through this process can only survive upto the blastocyst stage [24]. Parthenogenetic embryonic stem cell lines have been successfully derived from human blastocysts [25] in addition to non-human primates and mice. They behave like ES cells when cultured under appropriate conditions and are pluripotent, apart from contributing to chimera formation [23,26,27].

11.2.4 Germline-Derived Pluripotent Stem Cells

Germline-derived pluripotent stem (gPS) cells are usually derived from unipotent spermatogonial stem cells (SSCs) when established as germline stem cells (GSCs). These GSCs, when cultured under ES cell conditions, give rise to pluripotent stem cells, which are called germline-derived pluripotent stem (gPS) cells [28,29]. Recent studies reported the derivation of GSCs from testicular biopsies of 5–7-week-old adult mice. These techniques can eventually be used to obtain GSCs from human testicular biopsies, which will have great therapeutic potential [30,31].

11.2.5 Induced Pluripotent Stem Cells

Induced Pluripotent Stem (iPS) cells were first derived by Yamanaka et al. in 2006 from somatic cells by induction of pluripotency—a technology that has revolutionized stem cell research. Yamanaka and colleagues successfully reprogrammed embryonic or adult mouse fibroblasts by introduction of four factors: Oct4, Sox2 (SRY, or sex-determining region Y) box 2, c-Myc, and Klf4, using retroviral vector constructs (Figure 11.2). These cells are similar in all their properties to ES cells. They are pluripotent and have a capacity for self-renewal. They contribute to teratoma formation when injected into immunocompromised mice and contribute to the formation of all three germ layers [32]. Pluripotency can also be induced in adult human fibroblasts by addition of OCT4, SOX2, NANOG, and LIN28. These human iPS cells are similar to human ES cells in their ability to self-renew and become pluripotent [33]. They have also been derived from cells of several other animal species, such as pigs and bovine species, and potentially have great therapeutic value.

Induced pluripotent stem cells are derived by various methods. Reprogramming of somatic cells can be achieved by viral transduction (retroviral (lentiviral) and adenoviral). Retroviruses, including lentiviruses, pose the risk of insertional mutagenesis due to genomic integrations, while adenoviruses exhibit less efficient iPS cell generation and need repeated transductions. Recent studies report the use of Sendai viruses, which can be removed via antibody-mediated negative selection post-reprogramming, and are able to achieve efficient reprogramming [34,35]. Several nonviral methods of gene delivery have also been used for reprogramming, including nucleofection of nonviral polycistronic constructs and later excision of the reprogramming factors using a Lox/Cre system and the use of episomal vectors such as the *piggyback* transposon. These nonintegrating methods are usually less efficient for reprogramming. Indeed, most of the aforementioned methods are unsuitable for therapeutic purposes as they all involve the use of genetic material. For this reason, newer reprogramming

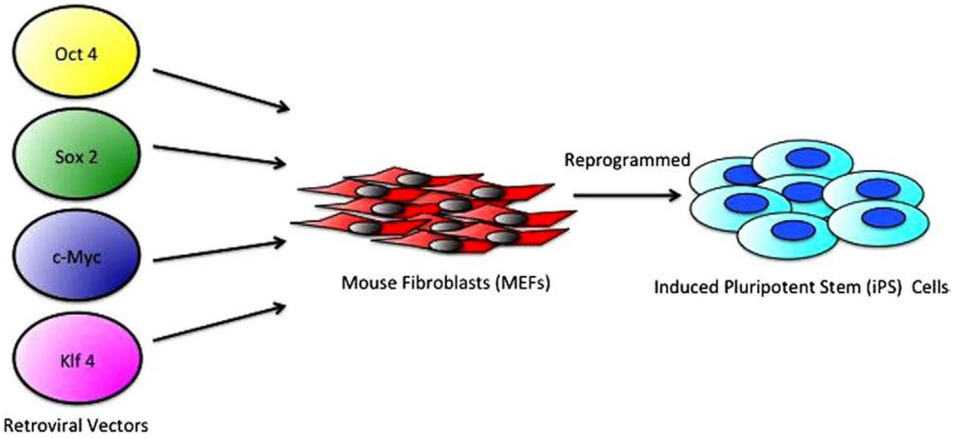


FIGURE 11.2 **Reprogramming of Mouse Embryonic Fibroblasts (MEFs).** The fibroblasts were successfully reprogrammed to pluripotency by overexpression of four factors—Oct4, Sox2, c-Myc, and Nanog—using retroviral vector constructs. The derived cells are similar to ES cells in all of their characteristics and are referred to as induced pluripotent stem (iPS) cells.

strategies that involve the use of mRNAs for delivery of reprogramming factors for iPS cell generation have also been described [34].

11.3 STEM CELL CULTURE METHODS

Cell culture methods are important for the physiological, pharmacological, and biological assessment of stem cells at tissue and molecular levels. Stem cells are usually cultured at 37 °C in the incubator with 5% CO_2 and 95% O_2. The composition of cell culture medium is important for maintenance of stem cells in an undifferentiated state. One of the most commonly used components in stem cell culture is fetal bovine serum (FBS). This, however, has several problems: batch-to-batch variability, the presence of undefined components such as hormones and growth factors (that may cause problems in cell culture) and so forth. Another major concern is the risk of transmission of proteins of animal origin to the patient. Scientists have thus moved to using serum-free culture conditions for making stem cell culture more relevant for clinical use [36].

Various culture systems are available for maintenance and storage of stem cells. For example, to separate MSCs from BM stroma, one of the proposed methods is the addition of 10 ml BM suspension to 20 ml of DMEM (Dulbecco's modified essential medium) containing 10% FBS; this solution can be centrifuged to separate the fat layer, and the cell pellet can then be subjected to density gradient centrifugation for 20 min at 13,000 g using a 70% Percoll solution. The MSC-enriched low fraction can further be cultured in DMEM, supplemented with 10% FBS, in a humidified incubator at 37 °C, and 5% CO_2 and 95% O_2. For further passaging, the cells can be enzymatically dispersed to single cells using 0.25% trypsin and 1 mM EDTA (ethylenediaminetetraacetic acid) and replated in DMEM supplemented with FBS for further maintenance [37]. For cryo-preservation, MSCs can be resuspended in 90% FBS and 10% DMSO (dimethylsulfoxide) by slow freezing and further keeping in liquid nitrogen for long-term storage [38]. For applications in regenerative medicine, MSCs can be can be isolated, stored, and cultured using xeno-free materials like serum-free MesenCult-XF™, TC-protector™, and recombinant trypsin (TrypLESelect™) [39].

For fetal stem cells, culture methods differ based on cell type. For example, cord blood can be collected from the neonatal umbilical cord of normal full-term pregnancies. It can then be sedimented with hydroxyethyl starch for 40–60 min. The mononucleocytes can be isolated by density gradient centrifugation at 400 g for 30 min, further washing in phosphate buffer saline, and resuspension in buffer solution. The CD34+ cells can be further isolated by positive selection using commercially available immunomagnetic separation systems. These isolated cells can be further counted and cultured in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% FBS [40]. For separation of MSCs from umbilical cord blood, the umbilical cord can be rinsed with phosphate buffer saline, disinfected in 75% ethanol, and further minced and digested. The cells can then be cultured in DMEM/F-12 medium supplemented with 10% FBS and 10 ng/ml epidermal growth factor (EGF), and passaged with trypsin. For cryo-preservation, the cells can be resuspended in the usual cryo-preservant solution, which consists of 90% FBS and 10% DMSO [41].

In the case of ES cells, several methods are available for maintenance, culture, and cryo-preservation. Those used for ES cell culture can usually be applied to the culture of iPS cells as well. Originally, human ES cells were derived and maintained in basal medium, which could be supplemented with fetal calf serum and grown on mouse embryonic fibroblast feeder layer (MEF). However, many factors in the serum are undefined and the quality of serum may vary across batches during production. Apart from this, the use of animal components in culture medium can pose a barrier for potential use of these cells for clinical applications.

To overcome these problems, researchers have developed more defined medium conditions for stem cell culture and maintenance. These conditions include hESF9 (consisting of basal medium supplemented with FGF-2 (fibroblast growth factor 2) and heparin sulfate), mTeSR1™, and STEMPRO™, which are commercially available and utilize several chemicals and growth factors to facilitate signaling in the promotion of human ES cell growth. The media are often supplemented with several components such as FGF, transferrin, albumin, cholesterol, and lipid mixtures [42]. ES cells can be passaged enzymatically, such as with dispase, accutase, collagenase, and trypsin, or mechanically, for example, STEMPRO® EZPassage™ disposable stem cell passaging tool, Life Technologies). Human ES cells are mostly frozen and thawed by the usual slow freezing and rapid thawing and subsequent addition of the ROCK inhibitor (Y-27632), which significantly improves recovery of the cryo-preserved cells [43]. The cells can be frozen in 60% KOSR (knockout serum replacement), 20% DMSO, and 20% DMEM in liquid nitrogen.

For parthenogenetic ES cells, the isolated ICMs can be cultured on mitomycin-C–treated mouse embryonic fibroblast cells (MEFs) for about 5–9 days; primary ES cell colonies can then be dissociated mechanically and cultured in human ES cell culture medium consisting of 80% knockout-DMEM, 20% serum-replacement, 1 mM glutamine, 1% nonessential amino acid, 0.1 mM 2-mercaptoethanal, 50 UI/ml penicillin, and 50 UI/ml streptomycin [25].

Human germ cells can be cultivated in gelatin-coated dishes in DMEM with high glucose, supplemented with 15% knockout serum replacement, 2 mM L-glutamine, 0.1 μM β-mercaptoethanol, 100 × nonessential amino acids, 1 × penicillin-streptomycin, 10 ng/ml bFGF (basic fibroblast growth factor), and 0.12 ng/ml TGF-β (transforming growth factor-β). These conditions lead to the formation of small ES cell-like colonies after four days of culture that are pluripotent [44]. The pES, gPS, and iPS cells can be cryo-preserved in ways similar to human ES cells.

11.4 MOLECULAR SIGNATURES OF STEM CELLS

As previously mentioned, stem cells are usually defined as cells that can both self-renew and differentiate into cells of various lineages. In the case of ES cells, the range of lineage options is large compared to other stem cell types such as hematopoietic stem cells, which are lineage restricted and can give rise only to blood cells. For this reason, several different stem cells types have certain shared properties as well as features that distinguish them from one another. It is therefore important to know how various cell fate decisions are made and how lineage commitment in stem cells is regulated. In the case of pluripotent stem cells such as ES cells, histone modifications, polycomb group (PcG) protein-binding patterns, and chromatin accessibility are important features that support pluripotency [45].

11.4.1 Epigenetic Modifications of Stem Cells

ES cells are generally derived from the ICM of pre-implantation blastocysts. These embryos are only transiently pluripotent, and pluripotency is reinstated later on by cells of the primordial germ layer. The primordial germ layer can also be used to generate other types of pluripotent stem cells such as embryonic germ (EG) cells as well as spermatogonial stem cells (SSCs) in adult males. When cultured pluripotent cells are reintroduced into a blastocyst, they give rise to chimeric animals whose adult tissues consist of stem cell–derived progeny. On the other hand, when ES or EG cells are injected into adults, they give rise to tumors known as teratocarcinomas, which consist of cells of all three germ layers. Also, ES and EG cells are capable of reprogramming somatic cells to enable them to re-express markers of their earlier embryonic stages. This requires that the expression of tissue-specific genes be switched off, whereas the expression of pluripotency-related genes such as Oct4 and Nanog must be switched on [45].

Cells of the early embryo have the ability to differentiate into all different cell types and are pluripotent. However, to differentiate into cells of other lineages, ES cells need to lose their potency, which is achieved by stable changes in their genome [46]. It is known that several heterochromatin and euchromatin structures, together with distinct DNA methylation and histone modification patterns, control gene activity and regulate chromatin state and gene expression. There are common epigenetic characteristic among different ES cell lines. Several mechanisms, termed epigenetic mechanisms, exist for regulating gene expression without changes in DNA sequences. The epigenetic information on a large number of genes in stem cells enables understanding pluripotency and mechanisms of genomic imprinting. It also helps in predicting the inclination of some ES cell lines to differentiate more readily than others. Thus, epigenetic signatures can serve as biomarkers for monitoring differentiation, self-renewal, and stem cell development [47]. These mechanisms are classified loosely into three categories: DNA methylation, histone modification, and gene regulation by small interfering RNAs [45].

11.4.1.1 DNA methylation

DNA methylation is a type of DNA modification during which a methyl group is added to the 5′ position of a cytosine base [48]. The process inhibits transcription mainly by two mechanisms: preventing the binding of the transcription factors to the DNA, or interacting with methyl CpG-binding proteins, which recruit histone-modifying and chromatin-remodeling factors. Methylation requires methyltransferases—for instance, DNMT1, which specifically recognizes the semimethylated DNA and goes on to methylate it. On the other hand, DNMT3A and DNMT3B can perform full demethylation [45,46]. DNA methylation and demethylation pathways are important for inducing differentiation in ES cells as well as for achieving pluripotency during somatic cell reprogramming, thus making it important for lineage specification as well as ES cell self-renewal [46,49].

DNA methylation involves the activation of two pathways. One, the AID (activation-induced cytidine deaminase) pathway has been speculated by some researchers to be involved in the demethylation of Oct-4 and Nanog promoters, which are master regulators of pluripotency. The second pathway enables the oxidation of 5-mC (5-methylcytosine) to 5-hmC (5-hydroxy-methylcytosine), 5-fC (5-formylcytosine), and 5-caC (5-carboxylcytosine) by TET (ten-eleven translocation) enzymes by base-excision repair. During somatic reprogramming, demethylation of pluripotency gene promoters can hinder the formation of fully reprogrammed iPS cells [46]. DNA methylation can be identified using several techniques. Most methods suitable for targeted as well as large-scale analysis use bisulphite treatment, which converts unmethylated cytosines to uracils with the methylated cytosine residues remaining unchanged. Bisulphite sequencing enables us to analyze DNA methylation at base-pair resolution. For genome-wide analysis, next-generation sequencing platforms can be used [48].

11.4.1.2 Histone Modifications

Histone molecules are made up of several amino acid residues; the N-terminal amino acids undergo various post-transcriptional modifications such as methylation, acetylation, phosphorylation, ubiquitination, sumoylation, citrullination, and ADP ribosylation. Histone modifications are associated with active as well as inactive chromatin sites. They are also associated with cellular processes such as mitosis, spermatogenesis, and DNA repair. In the case of certain modifications, specific proteins are recruited whereas in others the nucleosome structure is made more accessible to transcription factors through acetylation of residues. Thus histone modifications involve several synergistic as well as antagonistic interactions for the purpose of post-translational regulation [45].

Methylation of Lys27 on histone 3 (H3K27) is a major repressive histone modification and a hallmark of facultative heterochromatin. H3K27 methylation is catalyzed by enhancer of zeste homolog 2 (EZH2), which is part of the polycomb repressive complex 2 (PRC2). In embryonic stem (ES) cells, H3K27me3 is found at promoters of many important genes that regulate development. The mechanism of H3K27me3-mediated repression involves both PRC1 and PRC2. PRC1 acts via the H3K27me3 enzyme and catalyzes monoubiquitylation at Lys119 of H2A. This leads to activation of RNA polymerase II and represses the promoters. PRC1 also facilitates the compaction of chromatin, which is marked by H3K27me3, through a mechanism that is not dependent on its enzymatic activity. H3K27me3 also provides a mechanism to maintain and propagate H3K27me3-marked domains via PRC2. Loss of the repressive histone marks facilitates spontaneous ES cell differentiation. Thus, epigenetic marks prevent premature lineage commitment by silencing genes involved in differentiation.

However, the loss of PCR1 or PCR2 does not affect cell pluripotency [46]. Thus, H3K27me3 and H3K4me3 play a role in developmental gene expression. Loss of H3K27me3 may result in the deactivation of developmental genes that are normally not expressed in pluripotent cells. On the other hand, during lineage specification, the loss of H3K27me3 might result in both the deactivation of developmental control genes and misexpression of either pluripotency genes or other lineage genes. Loss of H3K4me3, on the other hand, may result in a less efficient induction of gene expression during differentiation, or it may result in loss or stochastic activation of gene expression [50]. H3 and H4 histones also undergo acetylation with the help of histone acetyltransferases, whereas deacetylation of histones is achieved by the action of HDAC (histone deacetylase) enzymes. Acetylation and deacetylation are both essential for proper ES cell differentiation as well as for maintenance of adult stem cell function. For example, the inhibition of HDACs prevents the differentiation of ES cells and is important for maintenance of their self-renewal [51].

Histone modifications can be analyzed by chromatin immunoprecipitation (ChIP) using antibodies that specifically recognize the histone tail modification of interest. Downstream methods for identifying exactly which DNA sequences are bound by histones with these modifications include targeted approaches such as quantitative real-time PCR (qPCR) and Sanger sequencing. High-throughput approaches to identify genome histone modifications include ChIP followed by microarray (ChIP-chip) analysis of predetermined regions (such as promoters) or ChIP followed by next-generation sequencing (ChIP-seq) for an unbiased analysis [48].

11.4.1.3 Small Interfering RNAs

Small interfering RNAs (siRNAs) are usually small antisense RNAs, approximately 20–25 nucleotides in length. They are generated from double-stranded RNAs and play a major role in RNA interference pathways by post-transcriptional regulation of gene expression [45]. These regulators of self-renewal in stem cells are part of a complex and interconnected transcriptional network. For this reason, the loss of function of one regulator can, in some cases, can be compensated for by adjusting the expression levels of other network components. Libraries can be developed for screening of siRNAs on a genome-wide scale [52].

11.4.2 Chromatin Modifications of Stem Cells

Chromatin refers to DNA and the proteins associated with it. The basic subunit of chromatin is the nucleosome, an octamer of four core histone proteins (two copies each of H2A, H2B, H3, and H4), around which ~147 base pairs of DNA are wrapped. Chromatin structure can be altered by histone modifications as well as by DNA methylation, which changes the DNA's density and accessibility, thereby modulating the transcriptional potential of the underlying DNA sequence. Histone variations and modifications can affect transcription either *in cis* (by sterically hindering DNA–protein interactions, by changing the charge of the chromatin, or by changing the stability of the nucleosome) or *in trans* (by creating binding platforms for downstream effectors).

Chromatin can also be remodeled and compacted by ATP-dependent chromatin remodelers recruited to specific locations in the genome by modified histones or by proteins with sequence specificity. Long-range interactions can also affect higher-order chromatin structure and transcription by bridging distant genome sites [50]. In ES cells, most genes involved in differentiation are regulated by chromatin-related mechanisms rather than DNA methylation [46]. Chromatin of undifferentiated ES cells is generally loose and less compact compared to differentiated cells, which makes the undifferentiated cells more available for transcription. It has also been shown that most of the transcription factors not expressed by stem cells, but essential for later developmental stages, replicate early in ES cells. Therefore, several lineage-inappropriate genes undergo changes in their chromatin structure and also switch to late replication, when ES cells are induced to differentiate. Gene promoters, which are highly conserved in ES cells, show a bivalent chromatin structure in which both active and repressive chromatin is closely placed. During differentiation, the bivalent chromatin profiles are generally lost and the tissue-specific genes undergo transcriptional activation, which further leads to silencing of the loci associated with alternative developmental pathways [45].

Apart from ES cells, the presence of bivalent chromatin structures is also seen in multipotent adult hematopoietic stem and progenitor cells, as well as neuronal progenitors. Chromatin-modifying complexes such as multiprotein repressive polycomb group (PcG) and activating Trithorax (TrxG) have long been known to regulate lineage specification during embryonic development [51].

11.4.2.1 Polycomb Group Protein Complexes

Polycomb group (PcG) proteins are also known to play a role in maintaining the inactive state of some developing embryos. PcG targets encode several transcription factors that are usually inactive in pluripotent cells but activated upon differentiation. PcG proteins may repress transcription in ES cells through a silencing mechanism that involves the formation of a complex chromatin structure. They may also alter the DNA topology by binding to DNA or to RNA. They are possibly recruited to specific sites in ES cells through stem cell-specific DNA binding factors. In these ways, PcG proteins play an important role in maintaining pluripotency [45]. In order to do so, they repress genes that regulate developmental pathways in such a way that the repression can be maintained epigenetically through cell division. Thus, they maintain the silent state of homeotic gene expression during embryonic development.

PcG proteins are composed of three distinct complexes: PRC1, PRC2, and PhoRC. PRC2 plays a role in the stabilization of repressive chromatin structures with the help of chromatin modifiers such as enhancer of zeste (EZH2), embryonic ectoderm development protein (EED), and suppressor of zeste 12 (SUZ12). All of these belong to the histone methyltransferases group and are responsible for depositing H3K27me2 and H3K27me3 marks onto chromatin. PCR2 is therefore a component of epigenetic memory, which is important for ES cell maintenance. PCR1, on the other hand, is useful in establishing a high-order chromatin structure. Through the activity of the E3 ligase enzyme, PCR1 is able to compact chromatin. Generally, PcG proteins can directly control DNA methylation and can act as connectors between key epigenetic events [51].

11.4.2.2 Trithorax-Group Protein Complexes

Trithorax-(TrxG-) group protein complexes mediate deposition of the histone post-transcriptional modifications responsible for active transcription [51]. They usually antagonize the function of PcG protein complexes, restricting their action to defined tissue-specific and development-specific sets of genes and excluding target genes from repression by PcG genes.

The TrxG protein complex mainly comprises three sets of proteins: the arabidopsis Trx homolog ATX1, the SAND-domain DNA binding protein ULTRAPETALA1 (ULT1), and the chromatin-remodeling protein PICKLE (PKL). ATX1 plays an important role in trimethylation of histone H3 at lysine 4 (H3Kme3). This modification is generally associated with regions that are transcriptionally active. It also regulates PcG target genes such as AGAMOUS (AG) and is involved in gene activation. SAND domains can bind to DNA and are involved in chromatin-associated transcriptional regulation. ULT1, on the other hand, is essential for full expression of AG. PKL is required for maintaining the transcriptional activity of target genes. The final outcome of all transcriptional activity at any given target gene is the result of the balance between the PcG and TxrG proteins together [53].

11.4.3 Alteration of Gene Expression in Stem Cells

Stem cells need to maintain normal expression of their genes, as alterations in gene expression can have a profound effect on cellular states and lead to pathological consequences. Stem cells give rise to progenitors, which are essential for maintenance of healthy organs. This is essential in case other cells in the organs are damaged or lost because of normal wear and tear or disease. During the process of aging, all organs lose cells [54]. The imprinting of maternal and paternal genetic components occurs very early during development. This involves the use of several epigenetic mechanisms. Noncompatible imprinting may lead to abortion or diseases such as malformation and cancer [54,55].

11.4.4 Genomic Imprinting in Stem Cells

Genomic imprinting is an epigenetic process that results in the preferential expression of monoalleles specific to the parent of origin. Most of the imprinted genes identified so far are physically linked in clusters and coordinate regulation. These clusters are very large and are highly conserved between mice and humans. Their structure is highly complex given that the same cluster usually comprises nonimprinted as well as maternally and paternally expressed genes. Imprinting centers (ICs) are the main cluster regions, which regulate the allele-specific expression of many genes [56–60].

11.5 EPIGENOMICS OF STEM CELLS

Epigenomics is defined as a genome-wide study of chromatin regulatory variation. This may include the study of histone post-transcriptional modifications, DNA methylation, nucleosome positioning, and long-range chromosomal interactions. Several methods are used to investigate the forms of this variation. These include chromatin immunoprecipitation (ChiP) for studying local histone post-transcriptional modifications at specific loci. For the analysis of DNA methylation patterns, bisulfite sequencing (BS-seq), restriction enzyme, and affinity approaches can be used. Similarly, genomic regions with low nucleosomal content can be identified by DNAse I hypersensitivity assays, whereas long-range chromosomal interactions can be identified by chromosomal conformation capture (3C) [61].

Recent breakthroughs in next-generation sequencing have allowed rapid adaptation and expansion of existing technologies for genome-wide analyses with better resolution and coverage. These methodologies include MethylC sequencing (MethylC-seq), which can be used for large-scale analysis of DNA methylation at single-nucleotide resolution. With the help of these modern techniques, chromatin signatures that characterize different classes of regulatory elements, such as promoters, enhancers, insulators, and long noncoding RNAs, can be unraveled.

ES cells can thus be used robustly *in vitro* to investigate the molecular and epigenetic basis for pluripotency and embryonic development. The chromatin of pluripotent cells has several unique features, including increased mobility of many structural proteins such as histones and heterochromatin protein 1. It also shares many fundamental properties with the chromatin of somatic cells. However, nuclear organization in ES cells is different from that in somatic cells and they have a less compact chromatin structure. The unique epigenome of ES cells promotes their epigenetic plasticity.

11.5.1 Epigenetic Variation among Pluripotent Stem Cells

Recent studies by Takahashi and Yamanaka reported the reprogramming of fibroblasts with forced introduction, via retroviral vectors, of the transcription factors Oct4, Sox2, Klf4, and c-Myc to generate iPS cells [32]. Induced iPS cells are successfully derived from a variety of somatic cells; they are phenotypically and molecularly similar to ES cells in many ways, including their gene expression patterns and their unique properties of pluripotency and self-renewal. However, the full extent of this similarity is still not clear. As with ES cells, tetraploid embryo complementation of most iPS cells can give rise to all tissues of the embryo proper. However, some iPS cell lines are unable to fulfill the criteria of pluripotency. Epigenetic studies comparing DNA methylation and histone modification patterns in ES and iPS cells may help unravel the differences in the behavior of these two cell types [61].

Most comparisons show that the majority of ES cell features are re-established during iPS cell reprogramming. However, subtle differences are observed in the DNA methylation and gene expression patterns between iPS and ES cells. They fall into three main categories: (1) experimental variability in cell line derivation and culture, (2) genetic variation among cell lines, and (3) systematic differences representing hotspots of aberrant epigenomic reprogramming [61].

The first category may not necessarily represent meaningful biological distinctions between the two cell types, but it can nevertheless be important when the quality or differentiation potential of the individual lines needs to be assessed. The second category is mainly a result of genetic variation between the different sources of the human cells or embryos from which these iPS cells or ES cells are derived.

Genetic variation is mainly due to the differences within cell lines with respect to DNA and histone modification patterns. This calls for using stringent statistical analyses when comparisons between ES and iPS cells are made [61].

Recent studies of specific chromatin features in lymphoblastoid cells indicate that human genetic variation has significant impact on epigenetic landscapes. They show that, irrespective of whether these cells are isolated from related or unrelated subjects, they exhibit individual as well as allele-specific, heritable differences in chromatin signatures. These differences can be largely attributed to underlying genetic variants [61–63]. Also, the differences in the epigenomes of ES and iPS cells may arise because of many other factors, such as maintenance of epigenetic memory (somatic memory) during reprogramming in iPS cells or selective pressure during reprogramming. However, understanding epigenetic differences between human ES and iPS cell lines is essential for harnessing their potential in regenerative medicine [61].

It has been observed that, in several defined models of mouse reprogramming, differences in gene expression and differentiation potential that are specific to cell type of origin exist, especially in early-passage iPS cells. Thus, the hypothesis that an epigenetic memory of previous fate persists in these cells is reinstated. This epigenetic memory is mainly due to the presence of residual somatic DNA methylation in iPS cells. Most of it is retained within regions in proximity to CpG islands. Because of this incomplete loss of somatic methylation, iPS cells are sometimes predisposed to differentiation into fates related to the cell type of origin and thus differentiation into other lineages can be restricted in them. However, continuous passaging, serial reprogramming, and further treatment with small molecules, which act as histone deacetylase inhibitors or as inhibitors of DNA methyltransferase activity, can erase this memory. Studies show that the remnants of somatic DNA methylation are transient and can be erased from iPS cells through replication and cell division [61].

A recent study also demonstrated that in spite of similar DNA methylation landscapes in iPS and ES cells, the presence of hundreds of differentially methylated regions (DMRs) can be observed. Moreover, only a small fraction of DMRs seem to fail in erasing somatic DNA methylation, but the vast majority of the differences in DNA methylation patterns mainly correspond either to hypomethylation (defects in the methylation of genomic regions marked in ES cells) or to the appearance of iPS cell-specific methylation patterns, which are generally not present in ES cells or in the somatic cell type of origin. These DMRs are mostly resistant to passaging, especially in late-passage iPS cells. This study was performed using a limited number of iPS and ES cell lines, genetic and experimental variation among individual lines may yet be found to be a big contributor to the reported DMRs. However, iPS cells of different genetic background and cell type of origin do share nonstochastic epigenomic hotspots that pertain to reprogramming [61,64].

Apart from erasing somatic epigenetic marks, it is essential that ES cell-like epigenomic features be reestablished during reprogramming. In iPS cells, in the areas surrounding the centromeres and telomeres, reduced DNA methylation may sometimes be observed. This biased reduction, which is observed in subcentromeric and subtelomeric regions, may be attributed to the blocks of H3K9me3 in these regions. Their surrounding aberrant DNA methylation patterns may play a role in rendering the telomeric and centromeric regions more resistant to epigenetic changes. Recent studies have observed the presence of histone variant H3.3 in the subtelomeric and subcentromeric regions of mouse ES cells and mouse embryo. This histone is generally involved in transcription-associated and replication-independent histone deposition. Apart from this, it is known that histone H3.3 is important in the maintenance of transcriptional memory during reprogramming by somatic cell nuclear transfer. For this reason, it can be speculated that a similar mechanism may contribute to the resistance of reprogramming in the regions surrounding the telomere and centromere in iPS cells [61].

All pluripotent cell types constantly anticipate alternative developmental fates and thus show the existence of poised promoters and enhancers in their epigenome. It has been clearly demonstrated that in iPS cell promoter regions, bivalent domains are reestablished. However, more studies need to be performed in order to assess the extent to which the chromatin signatures associated with these poised developmental enhancers are recapitulated in iPS cells. The existence of developmental enhancers, which are associated with genes that are inactive in ES cells but active in post-implantation steps of embryogenesis, suggests that proper enhancer reassembly to a ES cell-like state is essential for determining the differentiation potential of iPS cells. If the developmental enhancers are not properly epigenetically marked, an impaired differentiation of the iPS cells may result, although the transcriptional or promoter modification level in the undifferentiated state may remain unaffected. It is therefore important that the epigenomic profiling of enhancers and promoters be taken into consideration when determining the quality and differentiation potential of iPS cell lines [61].

11.5.2 Importance of Epigenomics in Pluripotent Stem Cells for Regenerative Medicine

Epigenomics can be useful in the study of human development and disease. Apart from this, it can be useful for translation of stem cell technologies into clinical applications. With the help of epigenomic profiling, annotation of functional genomic elements can be achieved in a genome-wide and cell type–specific manner. The distinct chromatin signatures obtained by epigenomics can help identify and distinguish between active and poised promoters and identify insulator elements and the presence of noncoding RNAs that may be transcribed in a given cell type.

As epigenetic marking of promoters and enhancers determines developmental potential, epigenomic mapping can be a better tool to identify the differentiation potential of iPS cells compared to that of transcriptome profiling. However, it is important to assess the quality of iPS cells and human ES cells before using them in translational applications. For this purpose, appropriate resources need to be developed. For example, although ChIP-seq analysis of chromatin signatures is extremely informative, relying on the quality of antibodies used, it requires the development of renewable, standardized reagents. Also, to assess the significance of epigenomic pattern variation, sufficient numbers of reference epigenomes need to be obtained from human ES and iPS cell lines. These should be representative of genetic variation and have been rigorously tested in a variety of differentiation assays [61].

Because of ethical and practical limitations, very little is known about the regulatory mechanisms that govern early human embryogenesis. However, human ES cell–based differentiation models offer a unique tool for isolation and study of cells that correspond to transient progenitor states arising during human development. Moreover, subsequent epigenomic profiling of human ES cells that have been differentiated *in vitro* along specific lineages can be used to define and identify the functional genomic regulatory space, or "regulatome," of any given cell lineage. This approach can be extremely useful for identification of genome-wide tissue-specific enhancers and silencers, which are highly variable among both different and closely related cell types.

Characterization of cell type–specific regulatomes can be useful for comparative analyses of gene expression profiles. In addition, through bioinformatic analysis of the underlying DNA sequence, they can be used for the prediction and identification of novel master regulators of specific cell fate decisions, which can subsequently serve as candidates in the development of direct trans-differentiation approaches. Moreover, mapping of enhancer profiles provides an enormous resource in the development of reporters for isolation and characterization of rare human cell populations, such as the progenitor cells that mostly arise only transiently during development. Ultimately, this knowledge will be helpful in the refinement of current differentiation protocols and derivation of a better-quality and well-defined population of cells that are safer and more appropriate for replacement therapies. Also, characterization of cell type–specific regulatomes will be useful for understanding noncoding variations in human disease [61].

During the past few years, genome-wide association studies have dramatically identified several new genetic variants associated with some of the most common human disorders, such as various types of cancer, type 2 diabetes, obesity, cardiovascular disease, Crohn's disease, and cleft lip/palate. However, most of these variants occur in the noncoding parts of the human genome, which suggests the existence of a large noncoding component in human phenotypic variation and disease. Several reports have also highlighted the critical role played by genetic aberrations, which occur within individual distal enhancer elements in human pathogenesis. So far, the exact role of regulatory sequence mutation in human disease has not been examined in detail, but, given the rapidly decreasing cost of high-throughput sequencing and multiple disease–oriented, whole-genome sequencing projects that are under way, the coming years will be crucial in allotting functional significance to disease-associated noncoding mutations.

Nevertheless, for this purpose it is important to identify and obtain cell types relevant to disease, as well as to characterize their specific regulatomes. In the future, the combination of pluripotent cell differentiation models and epigenomic profiling will be an important tool for uncovering the role of noncoding mutations in human disease. For example, if the disease of interest affects a particular cell type that can be derived *in vitro* from human ES cells, characterization of its reference regulatome may reduce the vast genomic regions implicated in a particular disease into a much smaller regulatory space that can be more effectively examined for recurrent variants associated with that disease.

The function of regulatory variants can be further studied using *in vitro* and *in vivo* iPS cell–based disease remodeling strategies. For example, disease-relevant cell types obtained from patient-derived and healthy donor–derived iPS cells can be used to study the effects of the disease genotype on regulatomes specific to a certain cell type. Moreover, given that most regulatory variants are not likely to be heterozygous in patients, loss or gain of chromatin features associated with those variants (such as p300 binding, histone modifications, and nucleosome occupancy) can be assayed independently for each allele within the same iPS cell line. Indeed, researchers have recently developed allele-specific sequencing assays, the results of which can be compared with allele-specific RNA-sequencing transcriptome analyses from the same cells. Such studies may provide valuable insights into the effects of disease-associated regulatory alleles on the transcription of genes located in relative proximity to the chromosomes [61].

Analysis of human ES cell and iPS cell chromatin landscapes have helped to provide important insights into the molecular basis of pluripotency, reprogramming, and early human development. Next-generation sequencing technologies such as ChiP and MethylC have been important in the study of the pluripotent cell epigenome. Several chromatin features, such as bivalent promoters, poised enhancers, and pervasive non-CG methylation are known to be more abundant in human ES cells as when compared with differentiated cell types. It is essential that future studies dissect the molecular function of these epigenomic attributes and their relevance to human ES cell biology.

Epigenomic tools have been widely used in the evaluation of iPS cell identity. In general, the epigenomes of the iPS cell lines are known to be very similar to those of human ES cell lines, although recent reports suggest that some differences in DNA methylation patterns are observed between the two. It will be important to understand the origins of these differences (that is, somatic memory, experimental variability, and genetic variation) and their impact on iPS cell differentiation potential, as well as to determine their clinical applications.

It is important that additional epigenetic features other than DNA methylation be thoroughly compared, including proper reestablishment of poised enhancer patterns. A more complete picture of the epigenomes of ES and iPS cells will be helpful for useful insight into the developmental decisions in humans, which will not only help to improve our understanding of human development but also facilitate the establishment of robust *in vitro* differentiation protocols for clinical applications. These advancements will in turn allow for generation of replacement cells for cellular transplantation and for development of appropriate *in vitro* disease remodeling strategies. With the help of *in vitro* disease models, epigenomic profiling can easily be achieved; this will promote understanding of the genetic basis of complex human disorders, especially where most of the causative variants are predicted to occur within the vast noncoding fraction of the human genome [61].

11.6 PHARMACOGENOMICS OF STEM CELLS

Cell-based assays are used to detect a probable genetic predisposition to toxicity during early drug development. Fewer reliable and stable sources of diverse cell populations are available for the study of genomics-based drug toxicity, but human pluripotent stem cells can be easily used for the development of toxicity assays for drug screening. Not only do these cells have a unique ability of self-renewal, but they are also genetically more stable than transformed cells and can differentiate into virtually any cell type in the body. Human embryonic stem cells and human induced pluripotent stem cells have similar differentiation potential, and their ability to differentiate into different cell types that are clinically relevant for disease remodeling, drug screening, and cell replacement therapy, has been studied extensively [65].

Many protocols have been published for directed differentiation of these cells into various cell types, such as cardiomyocytes, neurons and glia, pancreatic cells, and hepatocytes. However, these differentiated cells may not necessarily express all of the characteristics of mature adult cells. The liver plays a major role in drug metabolism *in vivo*. For this reason, development of *in vitro* protocols for differentiation of pluripotent stem cells into functional hepatocytes are of prime importance in toxicity assessment [65]. Hepatocytes have been generated by several researchers *in vitro*, using stepwise differentiation protocols based on sequential addition of several cytokines and growth factors. These protocols are based on the *in vivo* embryonic development of the mammalian liver, which includes initial differentiation into definitive endoderm, followed by hepatocyte specification and finally hepatocyte maturation. The purity of the cultures is further monitored at each stage of differentiation by checking for marker expression; markers that are typically expressed at that stage.

For purification and enrichment of cultures, several techniques are employed, including fluorescence-activated cell sorting (FACS) using antibodies for detection of surface markers (e.g., asialoglycoprotein receptor 1 and E-cadherin) and reporter assays for determining albumin or anti-trypsin expression. These techniques generate cell populations similar to hepatocytes. In this respect, they show an ability to store glycogen and to express albumin and cytochrome P450s, thereby also exhibiting functional similarity to hepatocytes [65–76].

Hepatocytes have a unique ability to metabolize drugs with the help of cytochrome P450 enzymes. Obtained from pluripotent stem cells, they show expression of these enzymes at levels lower than those of adult hepatocytes or adult liver. This can be a major setback in their use for toxicity assays and drug development [65].

11.6.1 Generation of Ethnically Diverse Pluripotent Stem Cells for Pharmacogenomic Studies

The derivation of functional hepatocytes able to metabolize drugs may help in the development of cell toxicity assays. Such assays can be useful for testing the pharmacogenomic effects of drugs before they are released in the market. Every human ES and iPS cell line has a unique genome, which promotes the possibility of obtaining a collection of pluripotent stem cells that represent the ethnic diversity of the population targeted by a particular drug. However, this is not the case in reality. Recent studies have shown that most human ES cells being used are from either European or East Asian ancestry and thus lack ethnic diversity. Ethical concerns as well as the technical expertise required for ES cell generation further contributes to their lack of diversity [65,77,78].

Problems surrounding the use of human ES cells can be easily overcome by generating human iPS cells from a variety of sources. Different human iPS cell types can be derived from many living individuals, allowing for the possibility of preselection of chosen genotypes to create collections of cell lines that serve a specific purpose. These may include cell lines that represent a specific genetic disease for research purposes. So far there have been no collections of human iPS cells that specifically represent variants of drug metabolism enzymes. However, there are plans for a collection of HLA variants, which may be useful in cell generation and, in the future, suitable for cell therapy in regenerative medicine [65].

11.6.2 Creating Stem Cell Banks for Pharmacogenomic Studies

There are challenges to be overcome in establishing the routine use of human iPS cells in pharmacogenomics studies. The first is the investment required for the collection of somatic cells from diverse individuals. Further on, infrastructural investments will need to be made to assess the quality of the collected cells. Assessment can be carried out using genomic analyses, which may help in cell characterization. Still, this challenge needs to be addressed before any cells are used in drug screening for pharmaceutical development [65].

Another challenge that needs to be overcome is developing human iPS cell lines which can be used for culture and differentiation to obtain the huge numbers of cells needed for toxicology and drug screening studies in large-scale analyses. Also, it is important that cells obtained from human iPS cells are able to represent the adult and fully functional form of the desired cell type in order to be considered for use in cell therapy and drug-screening strategies [65].

Genome editing in human stem cells can be successfully achieved with the help of meganucleases, zinc-finger nucleases, adeno-associated viruses, and transcription activator-like effector nucleases [79–84]. These allow efficient introduction of a variety of genetic alterations, ranging from single-nucleotide modifications to whole-gene addition or deletion, all with a high degree of target specificity into iPS cell lines. A strong rationale for using genome-edited cell models for phenotypic studies is that, by assessing the effect of a DNA variant on an isogenic background, potential confounders will be minimized [84]. This has the potential, in the near future, of replacing genetically modified mice. Genome editing will make it possible to go beyond disease modeling towards the discovery of therapeutics. For example, genome-editing tools should make it straightforward to insert reporters into the genomic loci of interest, allowing RNA-interference screens or small-molecule screens to identify genes and probes that have the desired functional effect [79–84].

11.6.3 Use of Stem Cells for Toxicogenomic Studies

The use of stem cells to assess the risk of toxic agent exposure to human beings can lead to new developments in toxicogenomics. The effect of several toxic reagents can be tested *in vitro* via pluripotent cell–based assay—for example, studies can be carried out on the epigenetic mechanisms that cause teratogenic effects in the embryo due to treatment with thalidomide, an otherwise non-mutagenic and non-cytotoxic compound. The effects of other non-mutagenic but tumor-promoting epigenetic toxicants such as DDT can also be tested *in vitro*. The role of stem cells in the development and progression of cancer can also be well studied [54,85,86].

With the help of embryonic stem cells, studies can be carried out on the effects of mutagens on undifferentiated embryonic cells as well as the differentiated somatic cells derived from them. Moreover, these cells serve as a tool to study the dose-dependent effect of mutagenic compounds. The mutagenic potential of DNA alkylating reagents such as *N*-ethyl-*N*-nitrosourea (ENU) and *N*-methyl-*N*-nitro-*N*-nitrosoguanidine (MNNG), as well as the DNA cross-linking and alkylating substances such as mitomycin C (MMC) can also be investigated on undifferentiated versus differentiated cells. There is a noted difference in the sensitivity to mutagens as well as rates of mutations between undifferentiated stem cells and their differentiated somatic derivatives [54,87].

Cellular responses differ. This is observed in stem cells with different genetic backgrounds and gender, where the genetic background of the test system is important in the interpretation of results. Similarly, cellular responses differ with respect to gender when mouse stromal stem cells are exposed to an extremely low-frequency EMF. For example, in colony-forming unit of fibroblasts (CFU-f) assay, it has been observed that the stromal cells from female mice show a reduction in proliferation while those from male mice do not [54,88,89].

11.7 POTENTIAL OF STEM CELLS

Stem cells can be used in a wide range of applications—particularly in the development and screening of new drugs as well as in future cell replacement therapies. Pluripotent stem cells can therefore serve as an important tool in detailed studies of various diseases. Such studies in turn can be useful in the development of new strategies for treatment of disease (Figure 11.3).

11.7.1 Use of Stem Cells for Disease Modeling

Stem cells offer immense opportunities in unraveling the pathological mechanisms of several human diseases. They can also be used to develop new therapeutic strategies in a preclinical setting. ES cells and the more recently developed iPS cell technology provide an unlimited source of cells unlike primary cell cultures. ES or iPS cell lines can be useful in extrapolating results from animal studies and in potentially overcoming some of the limitations of animal models. Recent studies have successfully shown that iPS

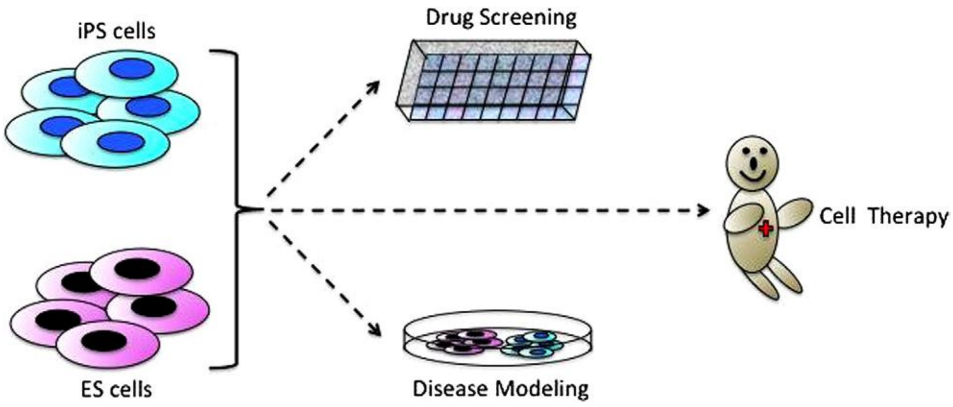


FIGURE 11.3 Potential of Pluripotent Stem Cells. Cells, such as iPS and ES, can be used for screening new drugs and modeling various human disorders *in vitro*; they also have potential in cell replacement therapy when differentiated into the desired cell type.

cells derived from patients may be useful for *in vitro* remodeling of inherited as well as multifactorial human disease. Thus they may in turn enable screening and validation of potential pharmacological compounds [90].

Robust protocols are available for differentiation of pluripotent stem cells into neurons; this has accelerated the development of stem cell–based models for the study of neurological disorders. Recently, iPS cells were derived from the fibroblasts of a child suffering from spinal muscular atrophy (SMA), which is an autosomal recessive genetic disorder caused by mutations in the SMN1 gene. Compared to the control iPS cell lines, the motor neurons derived from the SMA-iPS cell line showed a reduction in the expression of SMN1 protein, smaller size, and poor survival after prolonged culture *in vitro*. Notably, SMA motor neurons also responded to the SMN-inducing compounds, such as valproic acid and tobramycin, by upregulating SMN protein expression level [91]. This may provide a platform for future large-scale drug screening studies.

Recent studies have demonstrated that iPS cells can be used for modeling of diseases such as familial dysautonomia (FD), an autosomal recessive congenital neuropathy caused by a point mutation in the IKBKAP gene [92]. These studies have been successful in reproducing several aspects of the pathogenesis of the disease, including the mis-splicing of the IKBKAP gene and defective neurogenesis, which involve defective motility in FD-iPS cell–derived neural crest precursor cells. Long-term treatment with the drug kinetin seems to improve motility [92]. One study focused on X-linked MeCP2 gene mutations, which are commonly found in patients with the progressive neurodevelopmental disorder known as Rett syndrome. It reported that structural and functional abnormalities in neurons derived from Rett-iPS cells increased as a result of fewer dendritic spines, reduced synaptogenesis, abnormal calcium signaling, and electrophysiological defects [93]. Importantly, the beneficial effects of the previously identified drugs IGF1 and gentamicin in improving neuronal morphology and synaptogenesis [90] were confirmed.

The potential of iPS cells in modeling adult-onset and complex neurodegenerative diseases has been explored. In one study, iPS cell lines derived from fibroblasts in patients suffering from schizophrenia (SCZD) showed a decreased neuronal connectivity compared to healthy controls. In SCZD-iPS cell–derived neurons, the activity was improved by treatment with the antipsychotic drug loxapine [94]. In another study, iPS cells carrying a mutation in the LRRK2 gene, which is commonly associated with Parkinson's disease (PD), were differentiated into dopaminergic neurons and then exposed to stress agents. Compared to control iPS cell–derived dopaminergic neurons, PD-dopaminergic neurons showed an accumulation of a synuclein protein, an increased expression of oxidative stress response genes, and an enhanced sensitivity to neurotoxic reagents [95].

Very recently, iPS cells were derived from fibroblasts of patients with Alzheimer's disease (AD) in which mutations are present in the presenilin genes. Researchers observed that the neurons derived from AD-iPS cells had a higher secretion of amyloid β-42, which is a hallmark feature of familial Alzheimer's [96]. Moreover, these disease-specific iPS cells were responsive to g-secretase inhibitors, demonstrating the feasibility of this system for testing novel compounds for treatment of Alzheimer's [90].

Besides neurological disorders, iPS cell technology may prove useful as a tool for elucidating the molecular basis of cardiovascular and metabolic conditions, as well as for the development of assays for identification and screening of candidate drugs. For example, iPS cell–based modeling of long QT syndrome has been attempted independently by several groups and has demonstrated electrophysiological defects in long QT-iPS cell–derived cardiomyocytes. These defects can be partially restored by pharmacological agents [97–99]. Also, iPS cells derived from patients with LEOPARD syndrome have been shown to differentiate into cardiomyocytes that display hypertrophic features similar to those observed in patients [90,100].

Cell culture conditions for inducing definitive endoderm from pluripotent cells followed by hepatic specification are relatively well established. This has enabled several researchers to reproduce the key pathological events that define liver diseases *in vitro*. iPS cell lines have also been generated from patients suffering from inherited metabolic liver disorders. They have been successfully differentiated into hepatocytes or cells of the liver lineage. An aggregation of misfolded a1-antitrypsin (A1AT-alpha 1 anti-trypsin), for example, was observed in the endoplasmic reticulum of hepatocytes derived from iPS cells of patients with a1-antitrypsin deficiency [101]. Similarly, hepatocytes derived from iPS cells of patients with familial hypercholesterolemia (FH) had impaired LDL receptor–mediated cholesterol uptake, whereas those derived from glycogen storage disease type 1a patient fibroblasts showed accumulation of excessive amounts of lipid and glycogen and produced high levels of lactic acid [101]. As a first step toward an iPS cell–based model of type 1 diabetes, iPS cells were derived from fibroblasts that generated insulin-producing glucose-responsive cells *in vitro*. However, their efficiency was low [102]. Successful application of this model may require improved differentiation protocols as well as incorporation of immune cells isolated from the patient's blood or generated from iPS cells *in vitro* [90].

11.7.2 Use of Stem Cells for Toxicity Modeling

In recent years, immense improvement has been achieved in the development of experimental methods for screening compounds that are developmentally toxic and for elucidating the mechanisms responsible for their toxicity. Instead of animal models and highly laborious *in vivo* tests, we are now equipped with a rich source of human-derived materials, especially from ES and iPS cells. These show great potential in replacing conventional *in vivo* methods and are easily adaptable to a wide range of applications. Researchers are now focusing on completely validating and optimizing ES and iPS protocols that will significantly reduce or eliminate reliance on conventional and complicated *in vivo* protocols [103].

11.7.2.1 Use of Stem Cells in Cardiotoxicity Studies

It is important that the subclinical cardiotoxicity of new chemical entity (NCE) drug candidates is well assessed before their use in clinical applications. The detection of cardiotoxicity is crucial in preclinical stages, as it remains one of the leading causes of drug failure. Accurate and early detection of the effects of NCE drug toxicity is hampered, however, because of insensitive drug screening assays and assays that do not fully recapitulate the complex phenotype and exact effect of fully functional cardiomyocytes *in vivo* [103].

One of the most commonly observed cardiotoxic effects of NCE drugs is delayed repolarization of the ventricular action potential, mainly caused by transmembrane ion flux, which is detected electrophysiologically as a prolonged QT interval. A possible result is life-threatening ventricular tachyarrhythmia [104]. Most current toxicity tests are expensive and labor-intensive; they are encumbered by a high level of inter-experimental variability and by the large number of animals required. Also, the similarity between drug responses observed in animal models and in their human equivalents is not always observable because of huge amounts of interspecies variability.

To overcome these hurdles, mature cardiomyocytes isolated from humans are used as the cellular model for drug safety screening, although they are extremely difficult to harvest because of their limited availability, lower proliferative capacity, and disease state variability, especially when isolated from cadavers. Nevertheless, recent advances have created an unlimited source of cardiomyocytes that can be constantly replenished. Derived from human pluripotent stem cells, they can act as an alternative *in vitro* screening system. Much research is focused on the development of pluripotent stem cell–based cardiac models for testing toxicity. Early studies derived cardiomyocytes from mouse ES cells and hinted at their potential in drug safety screening and toxicology. However, because of interspecies variation, results based on mouse ES cells cannot be directly extrapolated to humans [103].

Several protocols have been developed for deriving cardiomyocytes from human ES and iPS cells. Some make use of spontaneous differentiation in three-dimensional spheroids known as EBs (embryoid bodies) [105]. Others use co-culture methods involving, for example, undifferentiated cells and END2 cells [106,107]. Generally, the number of beating EBs produced ranges from 0 to 70% of the total, and their composition is highly variable. Cardiomyocytes may make up only <25% of the total, with the rest mostly fibroblasts of endodermal lineage. Such huge heterogeneity may be due to differences in stem cell lines, in differentiation protocols, or in culture media, and methods used. Strategies to improve differentiation include the use of small molecules and growth factors. The efficiencies achieved are usually approximately 50%.

Selection can be used to obtain pure populations of cardiomyocytes from differentiating pluripotent stem cells. It can be carried out using manual methods such as dissection of contracting cells and centrifugation. An alternative is negative or positive selection for identified endogenous cardiomyocyte surface markers using fluorescent activated cell sorting. Cells can also be selected using reporter assays or marker genes. With any of these methods, cardiomyocytes with up to 99% purity can be obtained from differentiating pluripotent stem cell cultures [103].

Both of the widely used sources of pluripotent stem cells, iPS and ES cells, have been reported to differentiate into cardiomyocytes that have similar gene expression profiles and electrophysiological properties. Moreover, they are known to respond to cardioactive drugs in a similar way, although in some cases iPS cells may have impaired differentiation potential [108−110]. The differentiation of these cells into beating cardiomyocytes in terms of gene expression has been shown to be similar to the *in vivo* developing heart. Researchers have observed downregulation in the expression of pluripotency markers such as Oct4 and Nanog. They have also observed upregulation of cardiac transcription factors (e.g., NK2 transcription factor–related (NKX2.5) and myocyte enhancer factor 2C (MEFC2C)), as well as genes encoding structural proteins (e.g., myosin-heavy poly- peptide 7 (MYH7) and cardiac ion channels [111–113].

However, there are differences in pluripotent stem cell–derived cardiomyocytes compared to their adult phenotype. Immunofluorescence studies show the presence of cardiac structural proteins such as α-actinin and myosin light chain, but they exhibit relatively disorganized sarcomeric striations with a multi-angular appearance, similar to those of the human fetal cardiac phenotype, compared to adult cardiomyocytes, which have a more rod-shaped morphology and organized sarcomeric striations [106,114,115]. Also, pluripotent stem cell–derived cardiomyocytes lack the developed T tubule system that typically exists in mature cardiomyocytes [115]. Additionally, they behave electrophysiologically like 16-week-old fetal human ventricular cardiomyocytes and have a low resting membrane potential with low intracellular calcium handling [103].

Because pluripotent stem cell–derived cardiomyocytes can be maintained in culture for at least five months, their maturation can be induced if required [115,116]. It has been shown that the number and organization of myofibrils increase with time beyond day 50 of culture. In continuous culture, a high degree of sarcomeric organization may be observed in the differentiated cardiomyocytes with occasional formation of a T tubule system, indicating a phenotype that is compatible with mature cardiomyocytes [115,116]. The discrepancies between pluripotent stem cell–derived cardiomyocytes and their mature phenotypes may be due to secondary factors such as differences in the cell lines used, in the differentiation and culture protocols employed, and in the architecture of the cardiomyocytes being assessed. The relative immaturity of the differentiated cardiomyocytes ultimately may not matter as long as the cellular model is able to predict clinical outcome [103].

Microelectrode arrays (MEAs) and patch clamp techniques have been used to microscopically evaluate the beat alterations and electrophysiological changes, which are caused by several compounds, in pluripotent stem cell–derived cardiomyocytes. The results obtained have been as expected. The drugs evaluated include a variety of pharmacological classes such as adrenoceptor agonists and antagonists, muscarinic receptor agonists, acetylcholine inhibitors, calcium channel inhibitors, hERG potassium channel inhibitors, phosphodiesterase inhibitors, sodium channel inhibitors, and multiple ion channel blockers. Although derived *in vitro*, differentiated cardiomyocytes do show good predictivity of clinically observed cardiotoxicity, which makes them useful tools in drug discovery and safety pharmacology [103,117,118].

A complete assessment of the potential of stem cell–derived cardiomyocytes for use in drug toxicology can be carried out with the help of critical comparisons in sophisticated dose-response studies, using different concentrations of the drugs in patients under treatment. Direct comparisons with currently used *in vitro* and *in vivo* models are also required to demonstrate complete competency. To test the cardiotoxic effects of several known compounds, high- throughput screening platforms may be used for drug screening in 96-well formats using functional assays [103].

iPS cell–derived cardiomyocytes have several advantages over those derived from ES cells. For instance, they can be potentially used in safety pharmacology studies for selection of optimal medications for specific diseases. A recent study reported that an iPS cell–derived cardiomyocyte, carrying a congenital long QT syndrome type 2 mutation, developed early after depolarizations of the clinically used stressor isoprenaline. This effect was reversed by the beta-blocker nadolol, which has been used as therapy. This study concluded that iPS cell–derived cardiomyocytes respond appropriately to clinically relevant pharmacology, proving them to be a valuable *in vitro* model for disease modeling and individualized drug testing. From a drug-screening perspective, it may be possible in the future to generate a library of iPS cell–derived cardiomyocytes that reflect pharmacogenetic variations in the human population [103,119].

iPS cell–derived cardiomyocytes can also be used in unraveling the mechanisms associated with cardiac disorders. For example, in a recent study those from patients with the rare autosomal-dominant LEOPARD syndrome were used to study the pathophysiology of this disease *in vitro*. The obtained cardiomyocytes showed hypertrophy, which is a commonly encountered feature of patients in this group. Through their ability to model these genetically characterized diseases *in vitro*, patient-specific iPS cell–derived cardiomyocytes may also allow a deeper understanding of disease mechanisms and thus help to identify novel drug targets [100,103].

Pluripotent stem cell–derived cardiomyocytes show great potential in the evaluation of drug safety and toxicology studies. They are being increasingly used for disease modeling to understand the pathophysiology of drugs and to potentially identify new therapeutic targets. Though preliminary studies have been successful in drug safety evaluations

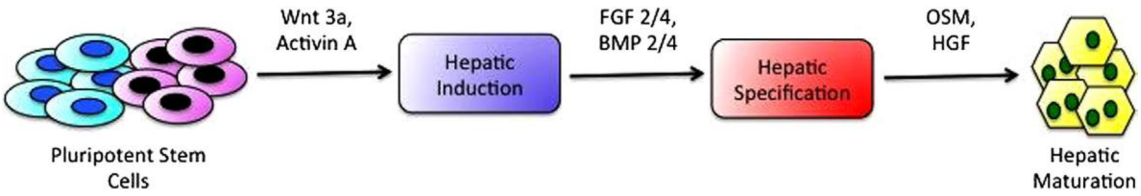


FIGURE 11.4 **Hepatic Differentiation of Pluripotent Stem Cells.** These cells can be differentiated *in vitro* in a stepwise manner by the addition of cytokines—Wnt 3a and Activin A help induce the cells toward hepatic differentiation. The differentiating cells can be specified toward a hepatic lineage by the addition of FGF 2/3 and BMP 2/4. Further maturation into hepatocytes can be achieved by the addition of OSM and HGF.

and high clinical predictions of cardiotoxicity, more studies are needed to confirm these results. Also, some practical issues need to be overcome. Recently reported methods are highly efficient, but newer differentiation and enrichment protocols need to be developed to routinely obtain pure cardiomyocyte populations. Also, these methods should be potentially usable in drug safety and drug toxicology evaluations in robust high-throughput screening formats, as well as scalable. Moreover, a refined cardiomyocyte population should be available for drug safety testing [103].

11.7.2.2 Use of Stem Cells in Hepatotoxicity Studies

Most of the current research in hepatic toxicology focuses on improving existing *in vitro* cellular models to optimize both their reliability in toxicology predictions and their ability to uncover associated metabolic pathways. The liver plays an important role in drug metabolism and excretion; adverse drug reactions (ADRs) are mostly caused by drug-induced liver injury [103,120].

The liver is a highly complex organ comprising several cell types. Usually preclinical drug testing for it is performed on several well-established *in vivo* and *in vitro* models from humans and rodents. However, interspecies differences in liver metabolism exist that make it difficult to extrapolate preclinical data to the clinic. For this reason, many drugs fail in human trials or are removed from the market because of hepatotoxicity. Efforts are being made to develop new models in order to optimize new compounds [103].

Since ES cells have been available for quite some time, a great deal of work has been done to obtain hepatocyte-like cells from them. Current protocols for the derivation of hepatocytes from ES cells usually focus on mimicking the pathways involved during *in vivo* liver development. These protocols can usually be categorized as definitive endoderm differentiation, hepatocyte progenitor specification, and hepatocyte maturation. Older techniques used EB-based protocols to imitate gastrulation for differentiation and lineage commitment. Because of the lower efficiency associated with this method, newer protocols mainly focus on direct ES cell stimulation to increase the specificity of lineage differentiation.

ES cell stimulation is usually facilitated by the addition of different exogenous factors (e.g., cytokines) to promote differentiation in accordance with different differentiation stages (Figure 11.4). For initial induction into the definitive endoderm, the ES cells can be treated with cytokines such as Wnt3a and Activin A. Rapid hepatocyte-like cell induction and maturation can be achieved by the addition of hepatocyte growth factor (HGF). Hepatocyte specification is obtained by addition of fibroblast growth factor (FGF 2/4) and bone morphogenic protein (BMP 2/4). For induction of maturation and to stimulate the cells to express metabolic enzymes associated with mature hepatocytes, Oncostatin M (OSM) and dexamethasone can be added. These factors are commonly found during hepatic development from the septum transversum and the cardiac mesoderm, and they are thought to induce expression of hepatic markers such as alpha-fetoprotein (AFP), albumin, cytokeratin-19 (CK19), HNF4α and HNF6 (hepatocyte nuclear factor), and A1AT. It has been found that small molecules can also be used to differentiate ES cells into hepatocytes [103,121–126]

In addition to exogenous factors, techniques have been employed to improve differentiation efficiency. For example, recent studies used co-culture with various cell types such as cardiac cells, non-parenchymal liver lineages, and feeder cells to replicate the *in vivo* conditions for hepatocyte differentiation. These cell types facilitate cell-to-cell interactions [127]. For the same purpose, some studies use dextracellular matrices such as three-dimensional collagen scaffold models [128]. Recent reports suggest using microRNAs. All of this work demonstrates a huge untapped potential to improve hepatic differentiation, especially through directed differentiation of pluripotent stem cells [103].

Transcriptome analysis shows similar metabolic and phenotypic characteristics in hepatocytes derived from ES and iPS cells. However, a difference in CYP isoform expression is observed in the hepatocytes derived from them. For example, hepatocytes obtained from ES cells show enriched CYP19A1, CYP1A1, and CYP11A1 expression, while iPS cell–derived hepatocytes show enriched CYP46A1 and CYP26A1 expression [129]. For this reason, it is important that, for any given differentiation experiment using either source, a full characterization of the resulting hepatocyte population be preformed, particularly by analysis of hepatic phenotype markers through gene expression analysis, functional assays, and activity analysis, as this defines the success of a hepatic differentiation experiment. The use of iPS cells for hepatocyte derivation *in vitro* is relatively new. However, as observed in ES cells, differentiation has been successfully achieved by several groups, using similar mechanisms with comparable efficiencies. Still, more work needs to be done to develop refined protocols for obtaining more accurate hepatocyte-like phenotypes in cells derived from both ES and iPS cells [103].

With the help of iPS cells, the possibility of modeling metabolic disorders *in vitro* can be explored. This may further aid in studies of hepatotoxicity screening. To validate the use of these cells for drug safety, metabolic studies, and clinical use, however, it is important to uncover the mechanisms underlying somatic cell–derived pluripotency.

The first iPS cells were developed through retroviral insertion of the pluripotency genes Oct4, Sox2, Klf4, and c-Myc, but generation efficiency was low and the possibility of genomic insertions due to oncogenes was high [32]. Since then, several approaches have been developed for inducing pluripotency by modifying and adapting the original technique using different methods of delivery, but these approaches are as yet limited by low efficiency rates, the use of viral vectors, the need for genome insertion, and highly complex protocols, and so they are not yet well optimized. This is an important issue because high-throughput screening of compounds requires high yields of genetically uncompromised pure cells [103]. To deal with it, alternative techniques, including the use of microRNAs, are being used.

The use of the miRNA cluster miR302/367 has been shown to induce pluripotency in embryonic mouse fibroblasts, with increasing efficiency as compared to the Yamanaka technique [32,103,130]. However, because a viral vector is still needed for delivery, it remains to be seen whether the use of microRNAs can be optimized with a view to scaling up production of human iPS cell–derived hepatocytes without viral vectors. MicroRNAs induce pluripotency in somatic cells through accelerated mesenchymal-to-epithelial transition, which causes demethylation of the genome; it also allows the derived cells to maintain an epigenetic memory because of incomplete reprogramming during transformation. Incomplete reprogramming commonly occurs during early derivation and in some cases may lead to improper differentiation into unrelated cell lineages. Epigenetic faults after differentiation may be associated in redifferentiated cell types from iPS cells. Although the mechanisms associated with this are still incompletely explained, these faults may be useful in improving the efficiency, quality, and purity of iPS cell differentiation [103].

With respect to disease modeling, a recent study has demonstrated that it is possible, in a parental hepatocyte with a metabolic disorder, to induce pluripotency and redifferentiate the cell into a hepatocyte-like cell that retains the diseased phenotype. This has been demonstrated in alpha-1-antitrypsin (A1AT) deficiency, familial hypercholesterolemia, glycogen storage disease type 1α, tyrosinemia, progressive familial hereditary cholestasis, and Crigler–Najjar syndrome. Thus, the possibility exists of modeling specific disease phenotypes for new therapeutic targets [101,131]. To explore the therapeutic potential of hepatocytes derived from iPS cells in liver diseases, zinc-finger nucleases (ZFN) and piggyBac can be used to correct a point mutation in the A1AT gene, thus causing restoration of function. This result highlights the clinical potential of iPS cells in future medicine, and it offers real hope for using pluripotent stem cells from different sources to produce hepatocyte-like cells that represent the phenotypic variations among different populations. For this purpose, it is important to examine the differentiated cells' ability to behave like primary hepatocytes in metabolizing drugs. Specific markers of differentiation need to be present, and the cells should be functionally similar to primary hepatocytes in their ability to secrete urea and albumin [103].

Hepatic metabolic enzymes can be split into two main groups: phase I and phase II. The most common phase I enzyme group is the CYP family, which has been associated with the formation of reactive metabolites and is vitally important for drug metabolism. Studies of CYP in pluripotent stem cells have produced varied results, and it has been suggested that the quality of the primary hepatocytes used as reference cells may be one of the causes of the large variations encountered. Much research is focused on the ability of CYP in hepatocyte-like cells to metabolize known compounds, with results suggesting detectable metabolic activity despite variations in results [103].

Very few studies have attempted to culture hepatocyte-like cells over a longer time period. This is an area in which improvement is needed to facilitate studies dedicated to uncovering effects of drug compounds over time. The risk is that the cells may dedifferentiate in long cultures. More studies are needed in order to fully understand the differentiation patterns of hepatocytes *in vitro* [103].

ADRs, particularly the idiosyncratic type, are often due to inter individual susceptibility to drug-induced liver injury. This is an important factor to consider when developing these model systems, as pluripotent stem cell–derived hepatocyte-like cells would have originated from one donor, thereby encapsulating genetic polymorphisms of that individual only. A panel or bank of hepatocyte-like cells from randomly selected donors address this, but further investigation will certainly require significant resources [103].

11.7.2.3 *Use of Stem Cells in Neurotoxicity Studies*

Modeling of neurotoxicity is very important, as damage to neuronal cells during development may result in lifelong learning disabilities and lead to mental retardation. For this reason, models need to be developed to facilitate the screening of new chemical entities (NCE) on developing and mature neuronal cells [103].

Most current models of neurotoxicity involve the use of immortalized or cancer cell lines, with extensive work being carried out in animals and primary animal cell cultures. However, it is difficult to obtain primary human material from the brain, as brain tissue does not regenerate, so fetal, instead of adult, tissue can be used. Still, fetal tissue raises ethical concerns and its availability is limited. Normal human neural progenitor cells are commercially available as a source of primary human fetal tissue, and they have been shown to successfully differentiate into functional and metabolically active neurons, astrocytes, and oligodendrocytes in culture. Before the development of protocols to differentiate stem cells into specific neuronal cell types, most neurotoxicity screening was usually carried out using immortalized cell lines taken from neuroblastoma and embryocarcinoma tissues, or derived from primary neural tissue using different techniques [103].

Significant progress has been made in the differentiation of stem cells to neuronal cells, particularly in neurotoxicity studies. Protocols have been developed in which human ES cells and human umbilical cord blood (UBC) stem cells, as well as human fetal tissue, can be differentiated into progenitor and mature neuronal cells. Differentiated neural progenitor cells can also be successfully maintained in culture for longer periods. Recent studies have reported further refinement of differentiation protocols with neurotrophic growth factors such as NT3 and NT4, and with maintenance of cell contacts [103].

Though some progress has been achieved, demonstration of toxicity with typical neurotoxins is still in its early stages. For example, the human umbilical cord blood neural stem cell line (HUCB-NSC) was tested for its ability to model neurotoxicity at three stages of differentiation: undifferentiated, committed (2 weeks in culture without RA treatment), and lineage-directed (2 weeks in culture with RA treatment). Here, compounds including six known neurotoxins and two non-neurotoxins were used and their toxicity effect on cells was analyzed, via live/dead viability/cytotoxicity fluorescence assay (staining for markers of metabolic activity and plasma membrane integrity); MTT assay; resazurin reduction cell viability assay; nuclear chromatin staining; and immunocytochemical labeling. These revealed markers of cell proliferation, apoptotic cell death, and neural differentiation. The HUBC-NSC model was found to be successful in discriminating between neurotoxic and non-neurotoxic compounds, and its responses to the compounds used were in accordance with published observations of the effects of these compounds on commonly used PC12 cells (a cell line derived from pheochromocytoma) and cerebellar granule cells. Researchers thus concluded that HUCB-NSCs are as sensitive as other current neurotoxicity models [103,132].

Several groups have differentiated human ES cells into neuronal cells for toxicity screening. The commonly used environmental neurotoxin methylmercury (MeHg) can be successfully employed to test the predictivity of several differentiated neuronal cell lines. MeHg is found in freshwater fish and so may cause toxicity through increased membrane lipoperoxidation. ES cells were differentiated to neuronal cells using a three-step differentiation protocol: plating derived progenitor cells on fibronectin in medium containing N-2 supplement and two-step maturation using human recombinant glial–derived neurotrophic factor. Toxicity was then tested using the resazurin cell viability assay and rtPCR for key neuronal cell markers such as nestin [103,133].

This study of MeHg was the first to show how and the degree to which MeHg affects the stages of differentiation. As an alternative, MEA (microelectrode array) can be used to assay spontaneous electrical activity in neurospheres derived from ES cells because, compared to cytotoxicity assays, it is highly sensitive to changes in function before cell perturbation. One problem with MEA is that it is plating-dependent and so batch-to-batch variation may be observed.

Other toxicity end points used include rtPCR for neuronal cell markers such as musashi. Cell lines such as hN2 are ES cell–derived and can be fully differentiated into neuronal cells for toxicity screening. hN2 is as predictive of neuronal toxicity as animal models. Several neurite outgrowth inhibitors such as lithium, orthovanadate, and brefeldin A have been successfully tested using it [103].

Stem cell-derived neuronal cells have also been tested for their ability to predict neurotoxicity in a range of compounds for which it has yet to be established. For example, neural-committed stem cells have been differentiated from random humans in a five-step protocol that uses the neural growth factors NEAA and N2. Differentiation to dopaminergic neurons has been successfully carried out using medium preconditioned on the PA6 stromal cell line. Drugs such as amiodarone have been assayed in high-throughput format using an ATP assay [103].

A major benefit of stem cell-derived neuronal cells is their ability to model neuronal dysfunction. A recent study used different human ES cell lines to obtain neuronal cells using mouse stromal P6 cells as feeders, in an attempt to model the features of Parkinson's disease (PD) with 1-methyl-4-phenylpyridinium (MPP+), a drug that depletes dopaminergic neurons and thereby mimics PD's disease process [103,134].

It can be seen that developments in the use of human stem cells as models for neurotoxicity are fast progressing. These developments involve starting cell types such as human ES cells, HUCB-NSC, and neural progenitor cell lines, and the use of newer model neurotoxicants. *In vitro* models are also being tested for their capacity to predict neurotoxicity of drugs whose risk to neuronal tissue has not yet been defined. Cell viability assays, which are routinely used as an end point of neurotoxicity, are still limited in their predictive capabilities. For this reason, combinations of viability assays and analysis of other markers of cell health (i.e., functional markers and markers of apoptosis and cell death) are more useful for monitoring the effects of neurotoxic compounds. Also useful are functional assays such as neurite outgrowth and MEA to test for spontaneous electrical activity [103].

11.7.3 Use of Stem Cells for Cancer Therapy

Cancer cells grow aggressively and proliferate continuously without a check. This uncontrolled growth results in invasion and sometimes complete destruction of adjacent tissues, which may then lead to patient death. Most of the time, cancer can be treated with surgery, radiation, and chemotherapy. However, these methods may not necessarily be curative and may often lead to high patient morbidity.

Nowadays, cell-based therapies are used to treat cancer. For example, human ES cell–derived cells can be successfully differentiated into NK cells, which produce cytokines and perform antibody-mediated or direct cell–mediated cytotoxicity on target cells. In this way, they can be used in cancer therapy, possibly making use of exogenously expressed chimeric antigen receptors or T-cell receptors to recognize cancer cell antigens and direct the cytotoxic cells to kill the cancer cells in response [135,136]. Dendritic cells, which present antigens, have also been derived from ES cells and have been used for targeting cancer cells.

For successful cancer treatment, genetic manipulations of ES cells are needed and further investigation is required. Recent studies have shown that the conditions and medium used for culture of human ES cells inhibits tumor growth. Also, it has been reported that the microenvironment of human ES cells neutralizes the expression of Nodal, an embryonic morphogen, in aggressive tumor cells and further inhibits tumor growth. These findings provide many new ways to use human ES cells in the treatment of cancer [136].

Despite their powerful differentiation capacity and relatively easily *ex vivo* culture characteristics, there are several key problems with human ES and iPS cells that must be overcome before they can be used in clinical applications. One such problem involves ethical concerns, which have limited research on this cell type. Even though human iPS cells overcome these ethical concerns, the efficiency of their generation still needs improvement. Also, they must be better characterized and their pluripotency established. In initial studies, both cell types were typically cultured on inactivated mouse embryonic fibroblasts (MEFs), which makes them inappropriate for clinical use because of xenogenic component contamination. However, improved culture systems use feeder-free conditions suitable for cell culture and maintenance.

It is difficult to achieve efficient differentiation of pluripotent stem cells because, in general, they induce teratoma formation. Their use in therapeutic applications thus remains a hurdle given that, within a population of differentiated stem cells, only a few undifferentiated cells are enough to form teratomas. Also, even though teratoma formation can be easily monitored, its incidence and mechanisms are not well understood, and few tools are available to eliminate it. Unless techniques are developed to completely shut down or control the pluripotency of undifferentiated stem cells in a population of differentiated cells, the undifferentiated cells will remain dangerous if used for therapy in humans. For this reason, much current research is focusing on developing better and more efficient methods for generating a homogenous and fully differentiated cell population derived from pluripotent stem cells.

Human ES- or iPS-derived cells in therapeutic treatment may still undergo immune rejection. If iPS cells are generated from and used in the same donor, this problem should be avoided, but it is still not well established whether, during the differentiation process, immunogenic molecules might be induced and expressed along the way. Such problems can be overcome by immune-suppressing drugs or genetic modifications of the pluripotent stem cells to improve immune tolerance; however, these solutions may have unwanted side effects.

Another way to overcome the problem of immune rejection is through cell banks of pluripotent stem cells based on HLA typing for matching these cells to patients. Of course, broad-based and large-scale information processing will be required to create such a useful database, but the safety of pluripotent stem cells in clinical applications will be improved. Also, banking will allow exploration of third-party cells for reprogramming, which may actually be a benefit since it is known that not all somatic cells can be reprogrammed with the same efficiency. Thus, cell sources that are abundant, easily available, and more efficiently reprogrammed can be used for human pluripotent stem cell generation, including fetal somatic cells. In this way, they can be further expanded and thus further extend the applicability of this revolutionary technology [136].

11.8 FUTURE PROSPECTS

The initial report on the reprogramming of somatic cells into iPS cells was followed by a remarkable upsurge in stem cell research. However, despite steady progress, crucial problems must be overcome for the field to realize its full potential. For one, differentiation protocols must be optimized to obtain reliable and safe cells. The increase in reprogramming efficiency was achieved without genetic cell modification, but the protocols used should be subject to rigorous quality control to maintain consistency and reproducibility so that reprogrammed cells can be effective in clinical settings. Also, genomic alterations that take place during reprogramming must be understood so that the derived cell populations will be pure, functional, and suitable for both research and therapy. To this end, it is important to minimize aberrations and generate more stringent pluripotency markers.

Refined differentiation protocols will help researchers investigate the pathophysiological basis of genetic diseases and carry out drug screening on affected cell types. Additionally, these protocols will facilitate derivation of patient-matched cells and tissues for clinical transplantation, which has been the main ambition of the stem-cell community and may be the ultimate measure of success in the field [137]. Moreover, because the use of iPS cells for drug development and testing requires patient-specific disease models, pluripotent stem cells may, in the future, aid the development of regenerative therapies by correcting cells with genetic disorders and further testing the functional capacity of these cells *in vivo* [138].

Pluripotent stem cells offer an intriguing tool for the study of mechanisms underlying various diseases as well as for the testing of various drugs against them. Pluripotent stem cell technology will eventually act as a boost to regenerative medicine by exploring various sources of unlimited cell populations to be used in clinical therapeutic settings.

ABBREVIATIONS

3C	chromosomal conformation capture
A1AT	alpha 1 anti-trypsin
AD	Alzheimer's disease
ADP	Adenosine Diphosphate
ADR	adverse drug reaction
AF	amniotic fluid
AFP	alphafetoprotein
AG	agamous
ATP	adenosine triphosphate
ATX	Ataxin
bFGF	basic fibroblast growth factor
BM	bone marrow
BMP	bone morphogenic protein
BS-seq	bisulphite sequencing
CD	cluster of differentiation
CFU-f	colony-forming unit of fibroblasts
CK	cytokeratin
CpG/CG	phosphodiester bond between cytosine guanine
CYP	cytochrome P
DDT	dichlorodiphenyltrichloroethane
DMEM	Dulbecco's modified essential medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
EBs	embryoid bodies
EDTA	ethylenediaminetetraacetic acid
EED	embryonic ectoderm development
EG	embryonic germ
EGF	epidermal growth factor
EMF	electromagnetic field
END2	endoderm-like cell 2
ENU	*N*-ethyl-*N*-nitrosourea
ES	embryonic stem
EZH2	enhancer of zeste
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
FD	familial dysautonomia
FGF	fibroblast growth factor
FH	familial hypercholesterolemia
gPS	germline-derived pluripotent stem
GSCs	germline stem cells
hERG	human *ether-à-go-go*–related gene
HGF	hepatocyte growth factor
HLA	human leukocyte antigen
HNF	hepatocyte nuclear factor
HSCs	hematopoietic stem cells
HUCB-NSC	human umbilical cord blood–neural stem cell line
IC	imprinting center
ICM	inner cell mass
IGF1	insulin-like growth factor 1
IKBKAP	inhibitor of kappa light polypeptide gene enhancer in B cells, kinase complex-associated protein
IMDM	Iscove's modified Dulbecco's medium
iPS/C	induced pluripotent stem/cell
Klf4	Kruppel-like factor 4
KOSR	knockout serum replacement
LDL	low-density lipoproteins
LEOPARD	lentigines, ECG anomalies, ocular hypertelorism, pulmonic stenosis, abnormal genitalia, retarded growth and sensorineural deafness
LRRK2	leucine-rich repeat kinase 2
MEA	microelectrode array
MeCP2	methyl CpG-binding protein 2
MEF/s	mouse embryonic fibroblast/s
MEFC2C	myocyte enhancer factor 2
MeHg	methylmercury
MethylC-seq	methylC sequencing
MMC	mitomycin C
MNNG	*N*-methyl-*N*-nitro- *N*-nitrosoguanidine
MPP	1-methyl-4-henylpyridinium
MSC	mesenchymal stem cell
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MYH7	myosin-heavy polypeptide 7
N-2	neutropan-2
NCE	new chemical entity
NEAA	nonessential amino acid
NKX2.5	NK2 Transcription Factor-Related
NT	neurotrophic
Oct4/OCT4	octamer-binding protein 4
OSM	oncostatin M
PC	pheochromocytoma
PcG	polycomb group
PD	Parkinson's disease
pES	parthenogenetic embryonic stem
PKL	Paxillin Kinase Linker
PRC	polycomb repressive complex
QT	Q wave T wave
RA	retinoic acid
RC	Repressive Complex
RNA	ribonucleic acid
ROCK	rho-associated protein kinase
rtPCR	real-time polymerase chain reaction

SAND	Sp100, AIRE-1, NucP41/75, DEAF-1
SCZD	schizophrenia
SMA	spinal muscular atrophy
SMN1	survival motor neuron 1
Sox2/SOX2	SRY (sex-determining region Y)-box 2
SSC	spermatogonial stem cell
SSEA4	stage-specific embryonic antigen 4
SUZ12	Suppressor of Zeste 12
TGF-β	transforming growth factor-β
TrxG	trithorax group
UCB	umbilical cord blood
UCMSC	umbilical cord matrix system cell
UCPVC	umbilical cord perivascular cell
UCSCs	umbilical cord stromal cells
ULT	ULTRAPETALA
ZFN	zinc-finger nuclease

REFERENCES

[1] Andersson ER, Lendahl U. Regenerative medicine: a 2009 overview. J Intern Med 2009;266(4):303–10.

[2] Mason C, Dunnill P. A brief definition of regenerative medicine. Regen Med 2008;3(1):1–5.

[3] Kumar R, Sharma A, Pattnaik AK, Varadwaj PK. Stem cells: an overview with respect to cardiovascular and renal disease. J Nat Sci Biol Med 2010;1(1):43–52.

[4] Mitalipov S, Wolf D. Totipotency, pluripotency and nuclear reprogramming. Adv Biochem Eng Biotechnol 2009;114:185–99.

[5] Kolios G, Moodley Y. Introduction to stem cells and regenerative medicine. Respiration 2013;85(1):3–10.

[6] Majo F, Rochat A, Nicolas M, Jaoude GA, Barrandon Y. Oligopotent stem cells are distributed throughout the mammalian ocular surface. Nature 2008;456(7219):250–4.

[7] Ousset M, Van Keymeulen A, Bouvencourt G, Sharma N, Achouri Y, Simons BD, et al. Multipotent and unipotent progenitors contribute to prostate postnatal development. Nat Cell Biol 2012;14(11):1131–8.

[8] Hipp J, Atala A. Sources of stem cells for regenerative medicine. Stem Cell Rev 2008;4(1):3–11.

[9] Biswas A, Hutchins R. Embryonic stem cells. Stem Cells Dev 2007;16(2):213–22.

[10] Spradling A, Drummond-Barbosa D, Kai T. Stem cells find their niche. Nature 2001;414(6859):98–104.

[11] Check E. Stem-cell paper corrected. Nature 2007;447(7146):763.

[12] Pappa KI, Anagnou NP. Novel sources of fetal stem cells: where do they fit on the developmental continuum? Regen Med 2009;4(3):423–33.

[13] Sanchez-Ramos J. Stem cells from umbilical cord blood. Semin Reprod Med 2006;24(5):358–69.

[14] Leitch HG, Blair K, Mansfield W, Ayetey H, Humphreys P, Nichols J, et al. Embryonic germ cells from mice and rats exhibit properties consistent with a generic pluripotent ground state. Development 2010;137(14):2279–87.

[15] Beddington RS, Robertson EJ. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development 1989;105(4):733–7.

[16] Bradley A, Evans M, Kaufman MH, Robertson E. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature 1984;309(5965):255–6.

[17] Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981;292(5819):154–6.

[18] Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci USA 1981;78(12):7634–8.

[19] Thomson JA, Kalishman J, Golos TG, et al. Isolation of a primate embryonic stem cell line. Proc Natl Acad Sci USA 1995;92(17):7844–8.

[20] Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282(5391):1145–7.

[21] Buehr M, Meek S, Blair K, Yang J, Ure J, Silva J, et al. Capture of authentic embryonic stem cells from rat blastocysts. Cell 2008;135(7):1287–98.

[22] Li P, Tong C, Mehrian-Shai R, Jia L, Wu N, Yan Y, et al. Germline competent embryonic stem cells derived from rat blastocysts. Cell 2008;135(7):1299–310.

[23] Kaufman MH, Robertson EJ, Handyside AH, Evans MJ. Establishment of pluripotential cell lines from haploid mouse embryos. J Embryol Exp Morphol 1983;73:249–61.

[24] Cibelli JB, Cunniff K, Vrana KE. Embryonic stem cells from parthenotes. Methods Enzymol 2006;418:117–35.

[25] Mai Q, Yu Y, Li T, Wang L, Chen MJ, Huang SZ, et al. Derivation of human embryonic stem cell lines from parthenogenetic blastocysts. Cell Res 2007;17(12):1008–19.

[26] Cibelli JB, Grant KA, Chapman KB, Cunniff K, Worst T, Green HL, et al. Parthenogenetic stem cells in nonhuman primates. Science 2002;295(5556):819.

[27] Allen ND, Barton SC, Hilton K, Norris ML, Surani MA. A functional analysis of imprinting in parthenogenetic embryonic stem cells. Development 1994;120(6):1473–82.

[28] Ko K, Tapia N, Wu G, Kim JB, Bravo MJ, Sasse P, et al. Induction of pluripotency in adult unipotent germline stem cells. Cell Stem Cell 2009;5(1):87–96.

[29] Kanatsu-Shinohara M, Inoue K, Lee J, Yoshimoto M, Ogonuki N, Miki H, et al. Generation of pluripotent stem cells from neonatal mouse testis. Cell 2004;119(7):1001–12.

[30] Ko K, Wu G, Arauzo-Bravo MJ, Kim J, Francine J, Greber B, et al. Autologous pluripotent stem cells generated from adult mouse testicular biopsy. Stem Cell Rev 2012;8(2):435–44.

[31] Ko K, Reinhardt P, Tapia N, Schneider RK, Araúzo-Bravo MJ, Han DW, et al. Brief report: evaluating the potential of putative pluripotent cells derived from human testis. Stem Cells 2011;29(8):1304–9.

[32] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126(4):663–76.

[33] Kastenberg ZJ, Odorico JS. Alternative sources of pluripotency: science, ethics, and stem cells. Transplant Rev (Orlando) 2008;22(3):215–22.

[34] Walia B, Satija N, Tripathi RP, Gangenahalli GU. Induced pluripotent stem cells: fundamentals and applications of the reprogramming process and its ramifications on regenerative medicine. Stem Cell Rev 2012;8(1):100–15.

[35] Li W, Zhou H, Abujarour R, Zhu S, Young Joo J, Lin T, et al. Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. Stem Cells 2009;27(12):2992–3000.

[36] de Sá Silva F, Almeida PN, Rettore JV, Maranduba CP, Maurmann de Souza C, Torres de Souza G, et al. Toward personalized cell therapies by using stem cells: seven relevant topics for safety and success in stem cell therapy. J Biomed Biotechnol 2012;2012:758102.

[37] Jaiswal N, Haynesworth SE, Caplan AI, Bruder SP. Osteogenic differentiation of purified, culture-expanded human mesenchymal stem cells *in vitro*. J Cell Biochem 1997;64(2):295–312.
[38] Mamidi MK, Nathan KG, Singh G, Trichelvam ST, Mohd Yusof NAM, Fakharuzi NA, et al. Comparative cellular and molecular analyses of pooled bone marrow multipotent mesenchymal stromal cells during continuous passaging and after successive cryopreservation. J Cell Biochem 2012;113(10):3153–64.
[39] Miwa H, Hashimoto Y, Tensho K, Wakitani S, Takagi M. Xeno-free proliferation of human bone marrow mesenchymal stem cells. Cytotechnology 2012;64(3):301–8.
[40] Yu X, Miao J, Xia W, Gu ZJ. Preferential expansion of umbilical cord blood-derived CD34-positive cells on human leukemia inhibitory factor transgenic feeder cells cultured on regenerated silk fibroin film. J Biomed Mater Res B Appl Biomater 2013.
[41] Gong W, Han Z, Zhao H, Wang Y, Wang J, Zhong J, et al. Banking human umbilical cord-derived mesenchymal stromal cells for clinical use. Cell Transplant 2012;21(1):207–16.
[42] Akopian V, Andrews PW, Beil S, Benvenisty N, Brehm J, Christie M, et al. Comparison of defined culture systems for feeder cell free propagation of human embryonic stem cells. *In Vitro* Cell Dev Biol Anim 2010;46(3–4):247–58.
[43] Rizzino A. Stimulating progress in regenerative medicine: improving the cloning and recovery of cryopreserved human pluripotent stem cells with ROCK inhibitors. Regen Med 2010;5(5):799–807.
[44] Golestaneh N, Kokkinaki M, Pant D, Jiang J, DeStefano D, Fernandez-Bueno C, et al. Pluripotent stem cells derived from adult human testes. Stem Cells Dev 2009;18(8):1115–26.
[45] Spivakov M, Fisher AG. Epigenetic signatures of stem-cell identity. Nat Rev Genet 2007;8(4):263–71.
[46] Li M, Liu GH, Izpisua Belmonte JC. Navigating the epigenetic landscape of pluripotent stem cells. Nat Rev Mol Cell Biol 2012;13(8):524–35.
[47] Deb KD, Jayaprakash AD, Sharma V, Totey S. Embryonic stem cells: from markers to market. Rejuvenation Res 2008;11(1):19–37.
[48] Lund RJ, Narva E, Lahesmaa R. Genetic and epigenetic stability of human pluripotent stem cells. Nat Rev Genet 2012;13(10):732–44.
[49] Fong YW, Cattoglio C, Yamaguchi T, Tjian R. Transcriptional regulation by coactivators in embryonic stem cells. Trends Cell Biol 2012;22(6):292–8.
[50] Vastenhouw NL, Schier AF. Bivalent histone modifications in early embryogenesis. Curr Opin Cell Biol 2012;24(3):374–86.
[51] Tollervey JR, Lunyak VV. Epigenetics: judge, jury and executioner of stem cell fate. Epigenetics: Official J DNA Methyl Soc 2012;7(8):823–40.
[52] Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, et al. Dissecting self-renewal in stem cells with RNA interference. Nature 2006;442(7102):533–8.
[53] Kohler C, Hennig L. Regulation of cell identity by plant Polycomb and trithorax group proteins. Curr Opin Genet Dev 2010;20(5):541–7.
[54] Ahuja YR, Vijayalakshmi V, Polasa K. Stem cell test: a practical tool in toxicogenomics. Toxicology 2007;231(1):1–10.
[55] Leclerc C, Webb SE, Daguzan C, Moreau M, Miller AL. Imaging patterns of calcium transients during neural induction in Xenopus laevis embryos. J Cell Sci 2000;113(Pt 19):3519–29.
[56] Rugg-Gunn PJ, Ferguson-Smith AC, Pedersen RA. Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines. Hum Mol Genet 2007;16(Spec No. 2):R243–51.
[57] Reik W, Walter J. Genomic imprinting: parental influence on the genome. Nat Rev Genet 2001;2(1):21–32.
[58] Smith RJ, Arnaud P, Kelsey G. Identification and properties of imprinted genes and their control elements. Cytogenet Genome Res 2004;105(2–4):335–45.
[59] Wood AJ, Oakey RJ. Genomic imprinting in mammals: emerging themes and established theories. PLoS Genet 2006;2(11):e147.
[60] Paulsen M, Takada S, Youngson NA, et al. Comparative sequence analysis of the imprinted Dlk1-Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2-H19 region. Genome Res 2001;11(12):2085–94.
[61] Rada-Iglesias A, Wysocka J. Epigenomics of human embryonic stem cells and induced pluripotent stem cells: insights into pluripotency and implications for disease. Genome Med 2011;3(6):36.
[62] Birney E, Lieb JD, Furey TS, Crawford GE, Iyer VR. Allele-specific and heritable chromatin signatures in humans. Hum Mol Genet 2010;19(R2):R204–9.
[63] McDaniell R, Lee BK, Song L, Zheng L, Boyle AP, Erdos MR, et al. Heritable individual-specific and allele-specific chromatin signatures in humans. Science 2010;328(5975):235–9.
[64] Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat Genet 2009;41(2):246–50.
[65] Fakunle ES, Loring JF. Ethnically diverse pluripotent stem cells for drug development. Trends Mol Med 2012;18(12):709–16.
[66] Agarwal S, Holton KL, Lanza R. Efficient differentiation of functional hepatocytes from human embryonic stem cells. Stem Cells 2008;26(5):1117–27.
[67] Basma H, Soto-Gutierrez A, Yannam GR, Liu L, Ito R, Yamamoto T, et al. Differentiation and transplantation of human embryonic stem cell-derived hepatocytes. Gastroenterology 2009;136(3):990–9.
[68] Cai J, Zhao Y, Liu Y, Ye F, Song Z, Qin H, et al. Directed differentiation of human embryonic stem cells into functional hepatic cells. Hepatology 2007;45(5):1229–39.
[69] Hay DC, Zhao D, Fletcher J, Hewitt ZA, McLean D, Urruticoechea-Uriquen A, et al. Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development *in vivo*. Stem Cells 2008;26(4):894–902.
[70] Hay DC, Zhao D, Ross A, Mandalam R, Lebkowski J, Cui W. Direct differentiation of human embryonic stem cells to hepatocyte-like cells exhibiting functional activities. Cloning Stem Cells 2007;9(1):51–62.
[71] Moore RN, Moghe PV. Expedited growth factor-mediated specification of human embryonic stem cells toward the hepatic lineage. Stem Cell Res 2009;3(1):51–62.
[72] Shiraki N, Umeda K, Sakashita N, Takeya M, Kume K, Kume S. Differentiation of mouse and human embryonic stem cells into hepatic lineages. Genes Cells 2008;13(7):731–46.
[73] Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, et al. Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. Hepatology 2010;51(1):297–305.
[74] Song Z, Cai J, Liu Y, Zhao D, Yong J, Duo S, et al. Efficient generation of hepatocyte-like cells from human induced pluripotent stem cells. Cell Res 2009;19(11):1233–42.
[75] Touboul T, Hannan NR, Corbineau S, Martinez A, Martinet C, Branchereau S, et al. Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. Hepatology 2010;51(5):1754–65.
[76] Liu H, Ye Z, Kim Y, Sharkis S, Jang YY. Generation of endoderm-derived human induced pluripotent stem cells from primary hepatocytes. Hepatology 2010;51(5):1810–9.

[77] Laurent LC, Nievergelt CM, Lynch C, Fakunle E, Harness JV, Schmidt U, et al. Restricted ethnic diversity in human embryonic stem cell lines. Nat Methods 2010;7(1):6–7.

[78] Mosher JT, Pemberton TJ, Harter K, Wang C, Buzbas EO, Dvorak P, et al. Lack of population diversity in commonly used human embryonic stem-cell lines. N Engl J Med 2010;362(2):183–5.

[79] Grizot S, Smith J, Daboussi F, Prieto J, Redondo P, Merino N, et al. Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. Nucleic Acids Res 2009;37(16):5405–19.

[80] Russell DW, Hirata RK. Human gene targeting by viral vectors. Nat Genet 1998;18(4):325–30.

[81] Khan IF, Hirata RK, Wang PR, Li Y, Kho J, Nelson A, et al. Engineering of human pluripotent stem cells by AAV-mediated gene targeting. Mol Ther 2010;18(6):1192–9.

[82] Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol 2007;25(11):1298–306.

[83] Hockemeyer D, Wang H, Kiani S, Lai CS, Qing G, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 2011;29(8):731–4.

[84] Musunuru K. Genome editing of human pluripotent stem cells to generate human cellular disease models. Dis Model Mech 2013;6(4):896–904.

[85] Davila JC, Cezar GG, Thiede M, Strom S, Miki T, Trosko J. Use and application of stem cells in toxicology. Toxicol Sci 2004;79(2):214–23.

[86] Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. Nat Rev Genet 2006;7(1):21–33.

[87] Cervantes RB, Stringer JR, Shao C, Tischfield JA, Stambrook PJ. Embryonic stem cells and somatic cells differ in mutation frequency and type. Proc Nat Acad Sci U S A 2002;99(6):3586–90.

[88] Klein TE, Altman RB. PharmGKB: the pharmacogenetics and pharmacogenomics knowledge base. Pharmacogenomics J 2004;4(1):1.

[89] Van Den Heuvel R, Leppens H, Nemethova G, Verschaeve L. Haemopoietic cell proliferation in murine bone marrow cells exposed to extreme low frequency (ELF) electromagnetic fields. Toxicol In Vitro 2001;15(4–5):351–5.

[90] Sommer CA, Mostoslavsky G. The evolving field of induced pluripotency: recent progress and future challenges. J Cell Physiol 2013;228(2):267–75.

[91] Ebert AD, Yu J, Rose Jr FF, Mattis VB, Lorson CL, Thomson JA, et al. Induced pluripotent stem cells from a spinal muscular atrophy patient. Nature 2009;457(7227):277–80.

[92] Lee G, Papapetrou EP, Kim H, Chambers SM, Tomishima MJ, Fasano CA, et al. Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. Nature 2009;461(7262):402–6.

[93] Marchetto MC, Carromeu C, Acab A, Yu D, Yeo GW, Mu Y, et al. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell 2010;143(4):527–39.

[94] Brennand KJ, Simone A, Jou J, Gelboin-Burkhart C, Tran N, Sangar S, et al. Modelling schizophrenia using human induced pluripotent stem cells. Nature 2011;473(7346):221–5.

[95] Nguyen HN, Byers B, Cord B, Shcheglovitov A, Byrne J, Gujar P, et al. LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. Cell Stem Cell 2011;8(3):267–80.

[96] Yagi T, Ito D, Okada Y, Akamatsu W, Nihei Y, Yoshizaki T, et al. Modeling familial Alzheimer's disease with induced pluripotent stem cells. Hum Mol Genet 2011;20(23):4530–9.

[97] Moretti A, Bellin M, Welling A, Jung CB, Lam JT, Bott-Flügel L, et al. Patient-specific induced pluripotent stem-cell models for long-QT syndrome. N Engl J Med 2010;363(15):1397–409.

[98] Itzhaki I, Maizels L, Huber I, Zwi-Dantsis L, Caspi O, Winterstern A, et al. Modelling the long QT syndrome with induced pluripotent stem cells. Nature 2011;471(7337):225–9.

[99] Yazawa M, Hsueh B, Jia X, Pasca AM, Bernstein JA, Hallmayer J, et al. Using induced pluripotent stem cells to investigate cardiac phenotypes in Timothy syndrome. Nature 2011;471(7337):230–4.

[100] Carvajal-Vergara X, Sevilla A, D'Souza SL, Ang YS, Schaniel C, Lee DF, et al. Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. Nature 2010;465(7299):808–12.

[101] Rashid ST, Corbineau S, Hannan N, Marciniak SJ, Miranda E, Alexander G, et al. Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. J Clin Investig 2010;120(9):3127–36.

[102] Maehr R, Chen S, Snitow M, Ludwig T, Yagasaki L, Goland R, et al. Generation of pluripotent stem cells from patients with type 1 diabetes. Proc Natl Acad Sci USA 2009;106(37):15768–73.

[103] Sison-Young RL, Kia R, Heslop J, Kelly L, Rowe C, Cross MJ, et al. Human pluripotent stem cells for modeling toxicity. Adv Pharmacol 2012;63:207–56.

[104] Pouton CW, Haynes JM. Embryonic stem cells as a source of models for drug discovery. Nat Rev Drug Discovery 2007;6(8):605–16.

[105] He JQ, Ma Y, Lee Y, Thomson JA, Kamp TJ. Human embryonic stem cells develop into multiple types of cardiac myocytes: action potential characterization. Circ Res 2003;93(1):32–9.

[106] Mummery CL, Ward D, Passier R. Differentiation of human embryonic stem cells to cardiomyocytes by coculture with endoderm in serum-free medium. Current Protoc Stem Cell Biol 2007. [chapter 1: unit 1F 2]

[107] Xu XQ, Graichen R, Soo SY, Balakrishnan T, Rahmat SN, Sieh S, et al. Chemically defined medium supporting cardiomyocyte differentiation of human embryonic stem cells. Differentiation 2008;76(9):958–70.

[108] Yokoo N, Baba S, Kaichi S, Niwa A, Mima T, Doi H, et al. The effects of cardioactive drugs on cardiomyocytes derived from human induced pluripotent stem cells. Biochem Biophys Res Commun 2009;387(3):482–8.

[109] Zhang J, Wilson GF, Soerens AG, Koonce CH, Yu J, Palecek SP, et al. Functional cardiomyocytes derived from human induced pluripotent stem cells. Circ Res 2009;104(4):e30–41.

[110] Yoshida Y, Yamanaka S. IPS cells: a source of cardiac regeneration. J Mol Cell Cardiol 2011;50(2):327–32.

[111] Anderson D, Self T, Mellor IR, Goh G, Hill SJ, Denning C. Transgenic enrichment of cardiomyocytes from human embryonic stem cells. Mol Ther 2007;15(11):2027–36.

[112] Beqqali A, Kloots J, Ward-van Oostwaard D, Mummery C, Passier R. Genome-wide transcriptional profiling of human embryonic stem cells differentiating to cardiomyocytes. Stem Cells 2006;24(8):1956–67.

[113] Honda M, Kiyokawa J, Tabo M, Inoue T. Electrophysiological characterization of cardiomyocytes derived from human induced pluripotent stem cells. J Pharmacol Sci 2011;117(3):149–59.

[114] Xu C, Police S, Rao N, Carpenter MK. Characterization and enrichment of cardiomyocytes derived from human embryonic stem cells. Circ Res 2002;91(6):501–8.

[115] Snir M, Kehat I, Gepstein A, Coleman R, Itskovitz-Eldor J, Livne E, et al. Assessment of the ultrastructural and proliferative properties of human embryonic stem cell-derived cardiomyocytes. Am J Physiol Heart Circ Physiol 2003;285(6):H2355–63.

[116] Baharvand H, Ashtiani SK, Valojerdi MR, Shahverdi A, Taee A, Sabour D. Establishment and *in vitro* differentiation of a new embryonic stem cell line from human blastocyst. Differentiation 2004;72(5):224–9.

[117] Dick E, Rajamohan D, Ronksley J, Denning C. Evaluating the utility of cardiomyocytes from human pluripotent stem cells for drug screening. Biochem Soc Trans 2010;38(4):1037–45.

[118] Laposa RR. Stem cells for drug screening. J Cardiovasc Pharmacol 2011;58(3):240–5.

[119] Matsa E, Rajamohan D, Dick E, Young L, Mellor I, Staniforth A, et al. Drug evaluation in cardiomyocytes derived from human induced pluripotent stem cells carrying a long QT syndrome type 2 mutation. Eur Heart J 2011;32(8):952–62.

[120] Baillie TA, Rettie AE. Role of biotransformation in drug-induced toxicity: influence of intra- and inter-species differences in drug metabolism. Drug Metab Pharmacokinet 2011;26(1):15–29.

[121] Baxter MA, Rowe C, Alder J, Harrison S, Hanley KP, Park BK, et al. Generating hepatic cell lineages from pluripotent stem cells for drug toxicity screening. Stem Cell Res 2010;5(1):4–22.

[122] Lavon N, Benvenisty N. Study of hepatocyte differentiation using embryonic stem cells. J Cell Biochem 2005;96(6):1193–202.

[123] Soto-Gutierrez A, Navarro-Alvarez N, Zhao D, et al. Differentiation of mouse embryonic stem cells to hepatocyte-like cells by co-culture with human liver nonparenchymal cell lines. Nat Protoc 2007;2(2):347–56.

[124] Tanimizu N, Miyajima A. Notch signaling controls hepatoblast differentiation by altering the expression of liver-enriched transcription factors. J Cell Sci 2004;117(Pt 15):3165–74.

[125] Duanmu Z, Locke D, Smigelski J, Wu W, Dahn MS, Falany CN, et al. Effects of dexamethasone on aryl (SULT1A1)- and hydroxysteroid (SULT2A1)-sulfotransferase gene expression in primary cultured human hepatocytes. Drug Metab Dispos 2002;30(9):997–1004.

[126] Kamiya A, Kinoshita T, Miyajima A. Oncostatin M and hepatocyte growth factor induce hepatic maturation via distinct signaling pathways. FEBS Lett 2001;492(1–2):90–4.

[127] Fair JH, Cairns BA, Lapaglia M, Wang J, Meyer AA, Kim H, et al. Induction of hepatic differentiation in embryonic stem cells by co-culture with embryonic cardiac mesoderm. Surgery 2003;134(2):189–96.

[128] Baharvand H, Hashemi SM, Kazemi Ashtiani S, Farrokhi A. Differentiation of human embryonic stem cells into hepatocytes in 2D and 3D culture systems *in vitro*. Int J Dev Biol 2006;50(7):645–52.

[129] Jozefczuk J, Prigione A, Chavez L, Adjaye J. Comparative analysis of human embryonic stem cell and induced pluripotent stem cell-derived hepatocyte-like cells reveals current drawbacks and possible strategies for improved differentiation. Stem Cells Dev 2011;20(7):1259–75.

[130] Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, et al. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell 2011;8(4):376–88.

[131] Ghodsizadeh A, Taei A, Totonchi M, Seifinejad A, Gourabi H, Pournasr B, et al. Generation of liver disease-specific induced pluripotent stem cells along with efficient differentiation to functional hepatocyte-like cells. Stem Cell Rev 2010;6(4):622–32.

[132] Buzanska L, Habich A, Jurga M, Sypecka J, Domanska-Janik K. Human cord blood-derived neural stem cell line–possible implementation in studying neurotoxicity. Toxicol In Vitro 2005;19(7):991–9.

[133] Zhang SC, Wernig M, Duncan ID, Brustle O, Thomson JA. *In vitro* differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol 2001;19(12):1129–33.

[134] Zeng X, Chen J, Deng X, Liu Y, Rao MS, Cadet JL, et al. An *in vitro* model of human dopaminergic neurons derived from embryonic stem cells: MPP+ toxicity and GDNF neuroprotection. Neuropsychopharmacology 2006;31(12):2708–15.

[135] Knorr DA, Kaufman DS. Pluripotent stem cell-derived natural killer cells for cancer therapy. Trans Res 2010;156(3):147–54.

[136] Ho PJ, Yen ML, Yet SF, Yen BL. Current applications of human pluripotent stem cells: possibilities and challenges. Cell Transplant 2012;21(5):801–14.

[137] Robinton DA, Daley GQ. The promise of induced pluripotent stem cells in research and therapy. Nature 2012;481(7381):295–305.

[138] Sharkis SJ, Jones RJ, Civin C, Jang YY. Pluripotent stem cell-based cancer therapy: promise and challenges. Sci Transl Med 2012;4(127):127ps9.

This page is intentionally left blank

Part IV

Statistics, Bioinformatics and Databases

This page is intentionally left blank

Chapter 12

Fundamentals of Complex Trait Genetics and Association Studies

Jahad Alghamdi and Sandosh Padmanabhan
Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK

Chapter Outline

12.1 Overview 235
12.2 The Phenotype 235
  12.2.1 Intermediate Phenotypes and Endophenotypes 236
  12.2.2 The Standardized Phenotype 237
  12.2.3 Assessment of the Genetic Basis of a Trait 237
  12.2.4 Familial Aggregation 237
  12.2.5 Recurrence Risk 237
  12.2.6 Heritability 237
12.3 Population Genetics 239
  12.3.1 Linkage Disequilibrium 241
  12.3.2 Linkage Studies 243
12.4 Association Studies 247
  12.4.1 Association Study Designs 247
  12.4.2 Population-Based Study Designs 249
  12.4.3 Correction for Multiple Testing 253
  12.4.4 Imputation 254
  12.4.5 Validation and Replication 255
  12.4.6 Meta-Analysis 255
  12.4.7 Fine Mapping and Resequencing 255
  12.4.8 Reporting of GWA Studies 256
12.5 Conclusions 256
References 256

12.1 OVERVIEW

There is significant inter-individual variability in drug response, much of which may have a genetic basis. This implies that genotype can influence drug metabolism, drug transport, and adverse drug reactions. A major aspect of pharmacogenomics involves using DNA sequence data to predict drug response and to inform drug discovery and development. This requires an understanding of population genetics and the methods available for gene mapping. In this chapter, we provide an outline of the principles of population genetics that underlie the analytic methods used in relating DNA sequence variation to drug response. We provide extracts from useful reviews that provide more detail on the concepts presented here.

12.2 THE PHENOTYPE

Phenotype or trait is the observable or measurable characteristic that is the target of genetic dissection. In pharmacogenomic studies, this can be measures of drug efficacy, side effects, or outcomes, and can be either quantitative or qualitative. The genetic architecture of a trait refers to all the genetic and environmental factors that contribute to the trait, as well as their magnitude and their interactions. The prevalence of the trait in the population is calculated as the phenotype frequency, which is the ratio of the number of individuals with a particular phenotype to the total number of individuals in the population. Phenotype frequency is a parameter that is usually calculated to determine whether the phenotype is common or rare (See Figure 12.1).

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00012-8
© 2014 Elsevier Inc. All rights reserved.

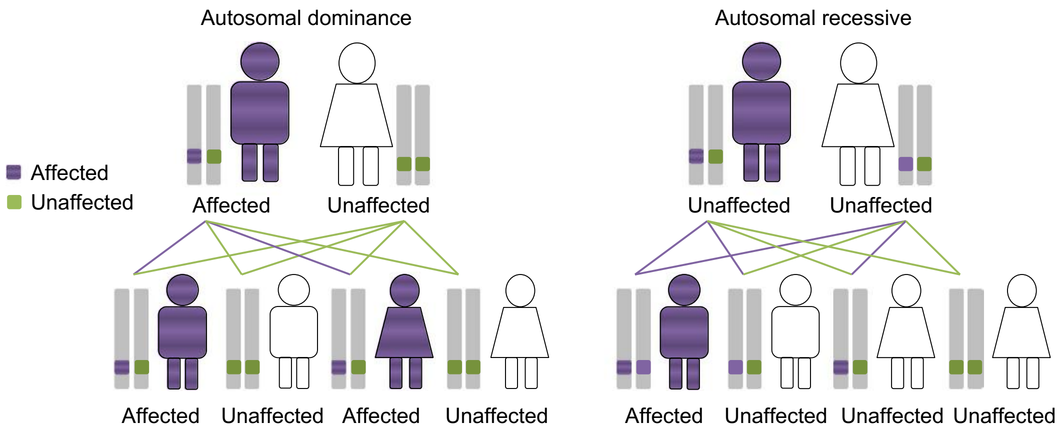


FIGURE 12.1 **Autosomal dominant and recessive inheritance.**

TABLE 12.1 **Simple versus Complex Traits**

	Simple	Complex
	• Single gene • Clear pattern of inheritance • Complete penetrance • No phenocopies	• Aggregate in families but do not segregate in Mendelian fashion • Multigenic • Multiple environmental factors • Phenotypic and Genetic heterogeneity • Phenocopies • Variable penetrance • Often not deleterious mutations, but natural variations frequent in the population • Complex interplay between genetic and environmental factors
Main problems	Genetic heterogeneity, pedigree size	Many unknowns about genetic basis: number of loci, number of alleles, allele frequency, risk or protection, interaction
Analysis	Parametric linkage analysis the method of choice: most powerful and not sensitive to allelic heterogeneity	Multitude of approaches with almost no ideal solution

A phenotype can be simple or complex and the characteristics of both are summarized in Table 12.1. Simple traits are rare, usually have a monogenic basis, and are due to a single gene with a large effect and high penetrance. Simple traits manifest Mendelian patterns of inheritance and can be expressed as dominant or recessive. Complex traits in contrast are more common, have a multifactorial etiology, and are due to numerous genes, environmental factors, and their interactions. Complex traits do not segregate within extended pedigrees in a straightforward Mendelian fashion.

Phenotypes can be dominant (when the trait is expressed in the presence of one copy of the allele) or recessive (when the trait is expressed only in homozygous form with two copies of the risk allele). The phenotype is co-dominant if both alleles are expressed in the heterozygous state.

12.2.1 Intermediate Phenotypes and Endophenotypes

In pharmacogenomics, the complexity of the processes that affect drug response and toxicity makes the selection of the right phenotype for study challenging. This is because a single disorder on a clinical level may actually be a heterogeneous group of potentially overlapping disorders on a genetic or etiologic level. All methods of gene localization, for example association studies, are sensitive to locus heterogeneity in which different genes contribute to phenotypic variation in different subsets of individuals. Association studies are also sensitive to allelic heterogeneity in which different variants in the same locus influence phenotype variation.

One strategy for dealing with underlying genetic heterogeneity, and a potentially complex network of etiologies contributing to phenotype variation, is to directly study quantitative risk factors that may index different underlying aspects of disease etiology. These are called intermediate phenotypes and they may be genetically less complex with potentially stronger genetic signals. The term "endophenotype" was coined in 1966 to distinguish between exophenotype (external) and endophenotype (internal) [1]. In general, an endophenotype must meet the following criteria: (1) it is linked to an illness in a population; (2) it is heritable; (3) it manifests itself whether or not the illness is active; and (4) it co-segregates with an illness in families [2].

It is generally easier to focus on a small group of genetic variants that have a large effect on the endophenotype. The closer a phenotype is to the level of gene action, the larger we might expect the genetic effects to be. For example, myotoxicity is the most common adverse effect of statins. Its pathophysiology is not fully understood, but it is likely to involve multiple pathways. Therefore, toxicity is defined in clinical trials by looking at a proxy, creatine kinase levels, which, if greater than tenfold the normal value, indicate toxicity; this is more tractable for genetic analysis [3]. Similarly, the readout for the toxic effect caused by inhibitors of the vascular endothelial growth factor (VSP) signaling pathway (e.g., sorafenib) is hypertension, so blood pressure can be used as an endophenotype for this toxicity [4].

12.2.2 The Standardized Phenotype

Accurate and standardized phenotyping is a vital initial step in genetic dissection of pharmacogenomic traits [5]. For example, given the relatively rare incidence of serious ADRs, only small numbers of well-phenotyped patients can be identified from a single center, necessitating the pooling of samples from multiple centers, often from different countries. A standardized phenotype thus becomes crucial for accurate pooling of data. Without a sufficiently large patient database, it would not be possible to identify genetic factors of modest effect or by ethnicity. Another critical point is that phenotypic misclassification may dilute or even mask a significant association with a gene locus, or it may provide a false-positive association. To help standardize phenotypic definitions, the Phenotype Standardization Project (PSP) was established to define in detail three serious forms of adverse drug reactions (drug-induced liver injury, drug-induced skin injury, and drug-induced *torsade de pointes*) that are used in GWAS [5]. The adoption of common phenotypic criteria is likely to benefit future research efforts, whether conducted by academic institutions, pharmaceutical companies, or regulatory bodies. The aim is creating a robust—and ultimately predictive—pharmacogenomic database that is relevant for all ethnicities [5].

12.2.3 Assessment of the Genetic Basis of a Trait

Familial aggregation analyses are traditionally used to obtain indirect evidence for a genetic contribution to a trait, by estimating the trait's heritable component. Heritability is the proportion of phenotypic variation due to the genetic differences between individuals within a population, and estimates of heritability can be used to describe the relative components of variance attributable to genetic and environmental factors [6,7].

12.2.4 Familial Aggregation

A shared genetic etiology for a trait means that family members are likely to share variants that influence the trait. In theory, phenotypes correlated with a particular risk allele should also be correlated among family members; this will lead to lower phenotypic heterogeneity [8]. Different metrics are used to quantify familial aggregation. If the phenotype is a binary trait coded for response or adverse effect, for example, the heritable component's recurrence risk ratio is calculated; on the other hand, if the trait is continuous, correlation or covariance-based measures such as the intra-family correlation coefficient (ICC) are used [6]. It is important to understand that assessment of familial aggregation takes no account of underlying biology. This means that evidence of familial aggregation does not always imply genetic effects. However, as most nongenetic factors associated with complex traits are weakly correlated in relatives and have only modest affects, the magnitude of familial aggregation can be considered to have a predominantly genetic basis.

12.2.5 Recurrence Risk

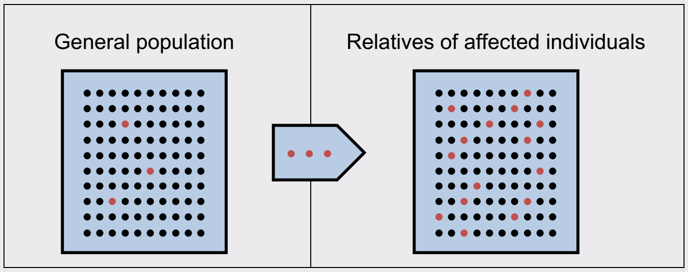
For a binary phenotype, the recurrence risk ratio (λ_R) is the ratio of disease prevalence in relatives (P_R) to the disease prevalence in the general population (P). λ_R can be estimated across different types of relatives; for example, λ_S (siblings) is the most common form. Measuring λ_R provides valuable information about the genetic component of the trait, and it is an important determinant of the power of affected relative pair studies to detect linkage (See Box 12.1).

12.2.6 Heritability

The heritability of a continuous trait is defined as the proportion of its total variance (σ_T^2) attributable to genetic

BOX 12.1 Calculation of Recurrence Risk

Disease prevalence (in percentage) is the number of individuals affected by a disease in a given population out of a total of 100 individuals at a particular time. Thus, 3 affected $=0.03$ and 15 affected $=0.15$. The recurrence risk for siblings (λ_S) is therefore $\lambda_S = 0.03/0.15 = 5$.



Examples of λ for Some Diseases

Type	Disease	P Relatives	P General	λ
Mendelian disorders	Cystic fibrosis	0.25	0.0004	500
	Huntington's disease	0.5	0.0001	5000
Complex disorders	Multiple sclerosis			~25
	Type 1 diabetes			~25
	Type 2 diabetes			~3.5

Complex diseases tend to have low λ, which reflects many genes, each contributing a small effect. Mapping is easier for traits with high λ, which suggests a major disease gene.

In late onset diseases, care should to taken that any systematic differences in the age distribution of the siblings of the affected cases and the the age distribution of the population sample from which population prevalence was estimated do not distort the statistic.

In relation to diseases with high population prevalence, such as asthma, λ_s has a theoretical upper bound at 1/prevalence, which means that one should be cautious in comparing values between populations that have different prevalences.

factors in a particular population. Narrow sense heritability (h^2) is defined as σ^2_A/σ^2_T, and broad sense heritability (H^2) is defined as σ^2_G/σ^2_T, where σ^2_A includes the additive genetic components of variance and σ^2_G includes all genetic components of variance [7]. For binary traits, heritability is usually calculated by invoking a hypothetical construct known as liability and applying a version of variance components modeling. Liability is an underlying, unobservable, normally distributed trait that is assumed to determine the probability that an individual will develop the disease of interest.

The main point to note here is that the numerator has a simple genetic meaning, but the denominator $-\sigma^2_T$ captures the variance attributable to genes and shared environment as well as residual variance attributable to unshared and unmeasured determinants and measurement error. Consequently, for a given trait heritability can vary quite substantially from study to study depending on the population being investigated as a result of varying environmental exposures between populations, the structure of the analytic model, and measurement error. If a disease process is entirely dependent on the presence of a particular allele of a particular gene, but all individuals in the population are homozygous for that allele, variation at that locus plays no role in variation in disease phenotype and therefore makes no contribution to heritability. On the other hand, the gene is clearly implicated in the causal architecture of the disease. Equivalently, a near ubiquitous environmental exposure makes little or no contribution to the denominator σ^2_T.

The power of most studies for discovering genes is positively associated with the heritability of the trait of interest. Thus, all else being equal and if the option exists, analytic efficiency may be enhanced by selecting a study population

TABLE 12.2 Heritability of Common Traits

Disease	Study Type	Heritability
High blood pressure	Twin studies	0.60
Asthma		0.72–0.8
Type 1 diabetes		0.72
Type 2 diabetes		0.26
Rheumatoid Arthritis		0.32
Body mass index	Family studies	0.34
Waist-to-hip ratio		0.28
Fasting glucose		0.33
Fasting insulin		0.23
Fasting triglyceride		0.47
Fasting high-density lipoprotein (HDL)		0.48
Systolic blood pressure (SBP)		0.30

in which heritability of the trait of interest is thought to be high. While the proportion of variance explained by genetic factors can provide insight into the value of genetic studies, it is important to note that heritability estimates do not provide information on the direction of effect of these factors or insight into their utility in prediction (See Table 12.2).

12.2.6.1 Heritability of Pharmacogenomic Traits

Studies in families can determine if common human disease phenotypes, like myocardial infarction or asthma have a heritable component. However, it is usually not possible to accumulate well-defined drug response phenotypes across multiple related patients with the same disease; as a result, the heritable component of variability in drug action may not be well-defined. One study used an *in vitro* approach where lymphoblastoid cells lines from related subjects were exposed to anticancer drugs, and the heritability of cytotoxicity was estimated at 25–65% [9,10]. Another approach is to quantify drug responses in multiple healthy members of a family—for example, twins. In twins, urinary excretion of isoniazid was found to be more variable in dizygotic twins compared to monozygotic twins; also, digoxin clearance was better correlated within monozygotic than dizygotic twins. Both findings indicate that shared genetic factors play a role in the expression of these drug related traits. The heritability of digoxin area under the curve was >79% [6]. In a study of the Amish—a founder population with extensive genealogical records—ADP-stimulated platelet aggregation showed a heritability of 33% at baseline and 73% after administration of clopidogrel, indicating a strong genetic component in clopidogrel drug response [11].

The heritabilities of some pharmacogenomic traits are presented in Table 12.3. If genome-wide genotype data are available for unrelated individuals, then there is the option to use linear mixed models to estimate the proportion of heritability explained by SNPs [12]. This can help determine the heritability of drug efficacy or toxicity especially from data in clinical trials.

12.3 POPULATION GENETICS

Population genetics is the study of Mendelian genetics at the level of a whole population. Genetic variants are classified according to their frequency or their nucleotide composition. Based on the minor allele frequency (MAF) in the population, a genetic variant is classified as common if its MAF is greater than 1%, or rare if its MAF is less than 1%. Common variants are also called polymorphisms (See Figure 12.2).

Single-nucleotide polymorphisms (SNPs) are the most prevalent class of genetic variation among individuals. It is estimated that the human genome contains ~11 million

TABLE 12.3 Crude Estimates of Heritability and Common Environmental Effects in Early Pharmacogenetic Studies in Twins

Drug	Pharmacogenetic Traits	r_{MZ}	r_{DZ}	h^2	C
Antipyrine	Plasma $t_{1/2}$	0.93	−0.03	0.93	0
Aspirin	Plasma concentration	0.9	0.33	0.85	0.05
Dicumarol	Plasma $t_{1/2}$	0.99	0.8	0.9	0.09
Ethanol	Absorption rate	0.56	0.27	0.4	0.16
Halothane	Percentage of dose	0.71	0.54	0.37	0.34
Isoniazid	Serum concentration	0.95	0.25	0.93	0.02
Lithium	Plasma concentration	0.94	0.61	0.85	0.09
Phenylbutazone	Plasma $t_{1/2}$	0.98	0.45	0.96	0.02
Salicylate	Steady-state level	0.64	0.32	0.47	0.17
Diphenylehydantoin	Plasma $t_{1/2}$	0.92	0.14	0.91	0.01

Source: Rahmioglu and Ahmadi (2010). Reprinted with permission from Future Medicine Ltd.
Note: $h^2 = (r_{MZ} - r_{DZ})/(1 - r_{MZ})$ and $C = r_{MZ} - h^2$. C: Common environment variance component; DZ: Dizygotic; h^2: heritabity, r_{MZ} and r_{DZ}: correlations within monozygotic and dizygoic twin pairs, respectively.

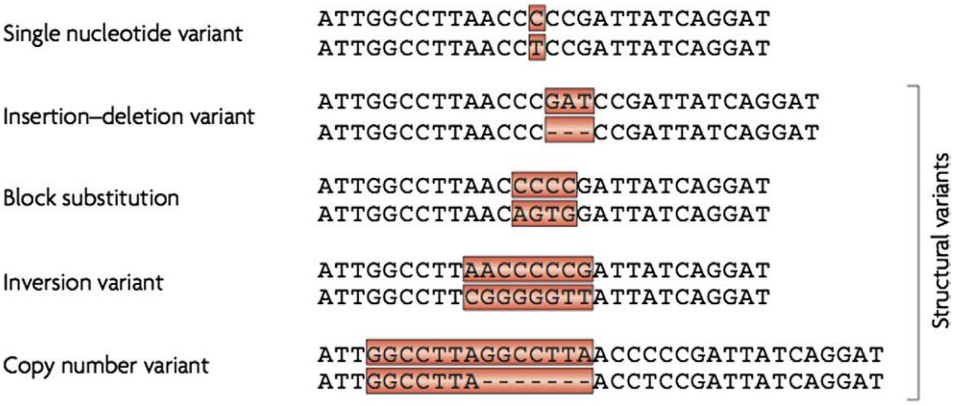

Single nucleotide variant
ATTGGCCTTAACCCCCGATTATCAGGAT
ATTGGCCTTAACCTCCGATTATCAGGAT
Insertion–deletion variant
ATTGGCCTTAACCCGATCCGATTATCAGGAT
ATTGGCCTTAACCC---CCGATTATCAGGAT
Block substitution
ATTGGCCTTAACCCCCGATTATCAGGAT
ATTGGCCTTAACAGTGGATTATCAGGAT
Inversion variant
ATTGGCCTTAACCCCCGATTATCAGGAT
ATTGGCCTTCGGGGGTTATTATCAGGAT
Copy number variant
ATTGGCCTTAGGCCTTAACCCCCGATTATCAGGAT
ATTGGCCTTA-------ACCTCCGATTATCAGGAT
Structural variants

FIGURE 12.2 **Types of genetic variations.** Single-nucleotide polymorphisms are a substitution of a single base pair in the DNA sequence. The insertion-deletion variants (indels) represent insertion or deletion of at least one base-pair in the DNA sequence. Block substitutions occur when a string of adjacent nucleotides varies between two genomes. Inversion variants describe the reversed order of nucleotides in a defined section. The copy number variants are repetitions of an identical sequence in some chromosomes. *Source: From Frazer et al.* [13]. *Reprinted with permission from Macmillan Publishers Ltd.*

SNPs, about 7 million of which have a MAF greater than 5%; the remaining have a MAF falling between 1% and 5%. Most of these SNPs are cataloged in the public dbSNP database and designated by a reference number. High-frequency SNPs are not expected to exert any major phenotypic effect because natural selection should ensure that SNPs are eliminated if harmful or fixed if beneficial. A genetic structural variant is any change in the human genome that involves more than a single nucleotide; such changes include insertion-deletion (indels), block substitutions, inversions of

DNA sequences, and copy number variants (CNVs) [13]. Four main factors make SNPs preferable to other types of genetic variations for dissecting the genetic basis of common diseases [14]. First, they are the most common and can be found throughout the genome, in exons, introns, promoters, enhancers, and intragenic regions, and some of them may themselves be functional. Second, when adjacent they show patterns of correlation that can enhance gene mapping and highlight recombination hotspots. Third, differences in their allele frequencies between different populations can be used in population-based genetic studies. And, fourth, they are more stable than other types of genetic variation, allowing more consistent estimates of genotype-phenotype associations.

12.3.1 Linkage Disequilibrium

Because the mutation rate is low (around 10^{-8} per site per generation) when set beside the most recent common ancestor of any two individuals (around 10^4 generations) [15], most SNPs are thought to arise from a single historical mutational event [14]. When a variant is first introduced into a population by mutation, it is perfectly correlated with nearby variants and is said to be in linkage. Linkage is the tendency of genes or other DNA sequences at specific loci to be inherited together as a consequence of their physical proximity on a single chromosome. Genetic recombination occurs when two homologous chromosomes exchange parts of their DNA during meiosis so that new combinations of alleles can be passed on to the next generation (see Chapter 2).

Recombination occurs at random. If there is a large distance between two segments of DNA, there is a good chance that recombination will occur between them. However, if the two sequences are close together, recombination will rarely occur between them—that is, the two sequences will tend to stay together rather than being split apart by recombination. In the latter case, the recombination frequency is low and the two DNA sequences are considered to be linked. High recombination frequencies are found for DNA sequences far enough apart that they are not linked to each other and therefore are considered to be in linkage equilibrium (i.e., both loci are segregating independently).

When a particular allele at one locus is found together on the same chromosome with a specific allele at a second locus—more often than expected if the loci were segregating independently in a population—the loci are in disequilibrium. Linkage disequilibrium (LD) is defined as the nonrandom association of alleles in adjacent loci, and following on from the preceding explanation, loci that are close to each other will be less likely to be split apart by recombination and thus be in linkage disequilibrium rather than equilibrium (Figure 12.3).

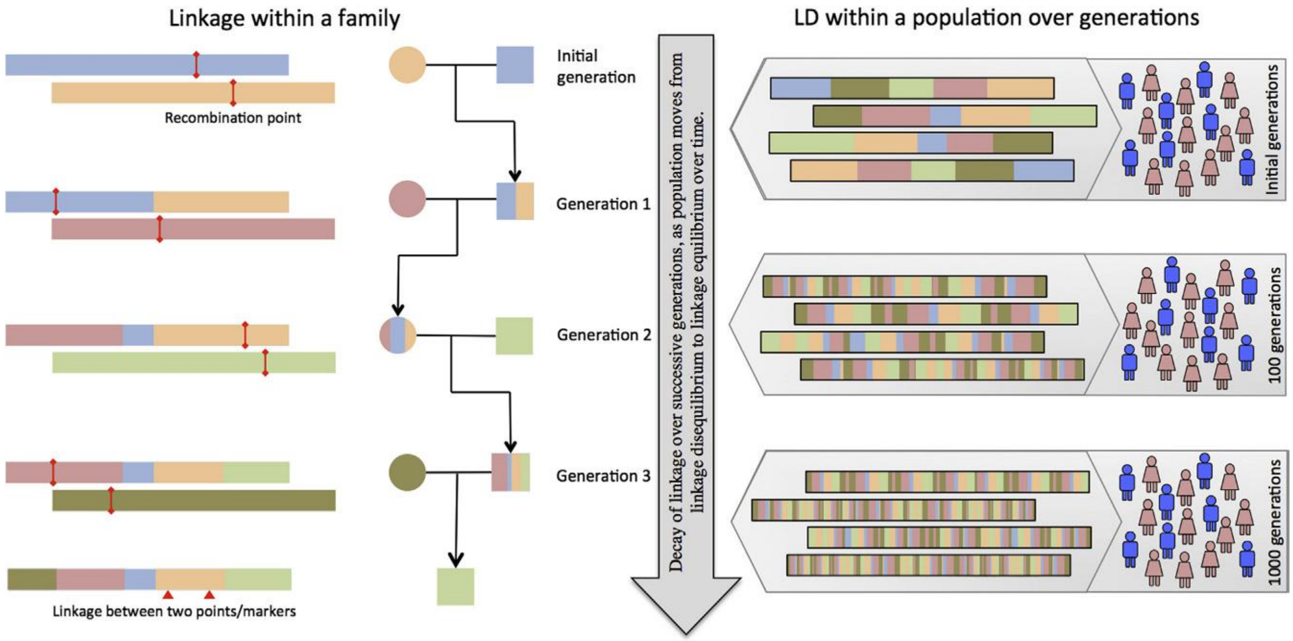

Linkage within a family
Recombination point
Initial generation
Generation 1
Generation 2
Generation 3
Linkage between two points/markers
Decay of linkage over successive generations, as population moves from linkage disequilibrium to linkage equilibrium over time.
LD within a population over generations
Initial generations
100 generations
1000 generations

FIGURE 12.3 Linkage and linkage disequilibrium. Within a family, a recombination occurs during meiosis in the germ cell between the paternal chromosome (blue) and the maternal chromosome (yellow). The offspring carry a recombinant chromosome, which undergoes further recombination in the next generation. In generation 3, the two markers (red triangles) in the ancestral (yellow) region are said to be in linkage. Within a population, the stretches from the ancestral chromosomes are decayed by recombination events over the time. At some point, a recombination event occurs between every possible point in the chromosome, which leads the population to move from LD to linkage equilibrium. *Reproduced with modifications Bush and Moore* [17].

When a variant is first introduced by mutation, it is in perfect LD (correlation) with its neighbouring loci. However, over successive generations this perfect correlation tends to break up (LD decay) due to a several reasons including recombination between the two loci, gene conversion, genetic drift, population growth, admixture, population structure, natural selection, variable recombination, and mutation rates [14,16].

A variety of statistics have been proposed to measure the amount of LD. The two most commonly used measures are D′ and r^2. The higher the value of D′, the lower the possibility that a recombination event occurred between these two loci. ($D' = 1$ means that the two markers have not been separated by a recombination event.) The absolute value of r^2 is more commonly used to quantify and compare LD in the context of mapping. When $r^2 = 1$, the two markers have not been separated by recombination and have the same allele frequency. In this case of perfect LD, the two markers are completely linked and observation at one marker provides complete prediction about the other. If the r^2 value between the causal marker and the tag SNP is less than 1, the sample size must be increased by $1/r^2$ of the sample size required for detecting the direct association with the disease's allele (See Box 12.2).

LD plays a fundamental role in gene mapping, both as a tool for fine mapping of complex disease genes and in genome-wide association studies. LD is also of interest for what it can reveal about evolution of populations. The rate of LD decay depends on population size, the number of founding chromosomes in the population, and the number of generations for which the population has existed [17]. For this reason, the pattern of LD differs among populations. For instance, African populations have smaller regions of LD because they are the most ancestral and so more recombination events have occurred. European and Asian populations have larger regions of LD because they were created by founder events that altered the number of founding chromosomes as well as population size, and generational age (See Box 12.3).

The international HapMap project has constructed genome-wide maps of LD patterns in multiple populations [15]. One of its main objectives is to identify set of SNPs that are in LD blocks to allow more efficient genotyping. These LD blocks vary in size but are typically 5–15 Kb and cover only 67–87% of the examined genomic regions. Interestingly, despite the great number of human DNA variants, human variability is much lower than expected, as each population has 4–5.6 haplotypes that account for 93–95% of copies of any given blocks. Thus, a small number of SNPs, called "tag SNPs", can be typed in each block to identify it (Figure 12.4). This implies that a limited set of SNPs (~500,000–1,000,000 SNPs) can capture ~90% of the genetic variation in the population [16]. Coverage is a measure of how well the SNPs that are part of a genotyping set capture all known variants; it is an important factor not only in the ability to capture all causal variants but also in terms of study power. The maximum r^2 (the squared correlation coefficient for each SNP) can be used to translate coverage to the sample size required for an indirect association study. For a particular variant, the effective sample size of such a study is simply the product of the actual sample size and the maximum r^2 value of the variant. Thus, as coverage decreases, a larger sample size is needed to obtain the same power. In turn, the overall power of a GWA study can be estimated using the effective sample size for each variant.

12.3.1.1 Hardy-Weinberg Equilibrium

In 1908, two scientists—Godfrey H. Hardy, an English mathematician, and Wilhelm Weinberg, a German physician—independently worked out a mathematical relationship that related genotypes to allele frequencies called the Hardy-Weinberg principle, a crucial concept in population genetics. It predicts how gene frequencies will be inherited from generation to generation given a specific set of assumptions. When a population meets all the Hardy-Weinberg conditions, it is said to be in Hardy-Weinberg

BOX 12.2 Applications of Linkage Disequilibrium Metrics

Several metrics have been devised to measure linkage disequilibrium (LD). The two most common are D′ and r^2. Both are related to D, the basic unit of LD.

D

D measures the deviation of haplotype frequencies from the equilibrium state. LD occurs when D is significantly greater than zero. Consider two linked SNPs with alleles (A, a) and (B, b), resulting in four possible haplotypes: AB, Ab, aB, and ab. D can be calculated as in equation 1, where f(X) represents the frequency of the X allele or haplotype.

$$D = f(AB) - f(A)\,f(B)$$

D′

D′ is the absolute ratio of D compared with its maximum value, D_{max}, when $D \geq 0$ or when compared with its minimal value,

D_{min}, when $D < 0$. $D' = 1$ denotes complete LD, and historical recombination results in its decay of toward zero.

r^2

r^2 is the statistical coefficient of determination—it measures correlation between a pair of variables.

$$r^2 = \frac{D^2}{f(A)\,f(a)\,f(B)\,f(b)}$$

r^2 is of particular importance in genetic mapping because it is inversely related to the required sample size for association mapping, given a fixed genetic effect. For example, if only one of a pair of SNPs is genotyped and r^2 between the SNPs is found to be 0.5, then, to provide the same statistical power for the ungenotyped SNP compared with the case where $r^2 = 1$, the sample size needs to be doubled. When $r^2 = 1$, knowing the allele genotypes of one SNP is directly predictive of the genotypes of another SNP. The alternative notation R^2 is used when individual variables are predicted using the multiple regression of a constellation of other variables.

Relationship between D' and r^2

D' and r^2 can be written in terms of each other and in terms of allele frequencies. Without losing generality, the four alleles can be chosen such that $D \geq 0$ and $f(A) \geq fB$. Thus, D' and D_{max} have the relations.

$$D' = \frac{D}{D_{max}}$$

$$D_{max} = f(a)f(b)$$

and

$$r^2 = (D')^2 \times \frac{f(a)f(B)}{f(A)f(b)}$$

This equation shows the relationship between D' and r^2 and allele frequencies. $f(A) \geq fB$, r^2 has the upper bound of $(D')^2$ and reaches it only when $f(A) = f(B)$, which implies that D', a commonly used measure of historical recombination, provides information on the physical extent of useful LD (in terms of association mapping and statistical power) by providing the upper limit of r^2. Dense LD maps based on high-frequency SNPs ($MAF > 0.1$) can reveal regions of historical recombination. Knowing the level of D' decay in these maps directly provides the maximum potential level of useful LD in association mapping (based on r^2) for high-frequency SNPs, even if a significant number of common SNPs remain undiscovered. For example, if a recombination point results in a D' of 0.7 for SNPs on either side of it, the maximum possible r^2 for these SNPs is 0.49, and sample sizes need to be more than doubled to maintain the same statistical power for association mapping. It should be noted that both D' and r^2 suffer from sampling biases given a small number of individuals and rare variants. Confidence intervals for D' have been used by some investigators.

Source: *Wang et al. [23]. Reprinted with permission from Macmillan Publishers Ltd.*

equilibrium (HWE). Human populations do not meet all the conditions of HWE exactly, and their allele frequencies will change from one generation to the next, so the population evolves. How far a population deviates from HWE can be measured using the "goodness-of-fit" or chi-squared test (χ^2) (See Box 12.4).

12.3.1.2 *HWE in Association Studies*

When marker allele frequencies in controls deviate markedly from HWE, this is commonly taken as an indicator that the genotyping is unreliable [18] or that there is marked population stratification (see the following). A slight departure from HWE may also be indicative of selection, in which case there will be deviation from HWE at loci associated with disease. Removing such loci should be carefully considered. The threshold of HWE should be carefully considered as well, particularly in the setting of GWA studies where a large number of SNPs are expected to deviate from HWE only by chance. For example, a GWA study that genotypes 500,000 SNPs may, by chance, have 25,000 SNPs with a p-value <0.05 for HWE. Therefore, a stricter threshold is required for GWA studies to identify worrisome HWE deviation.

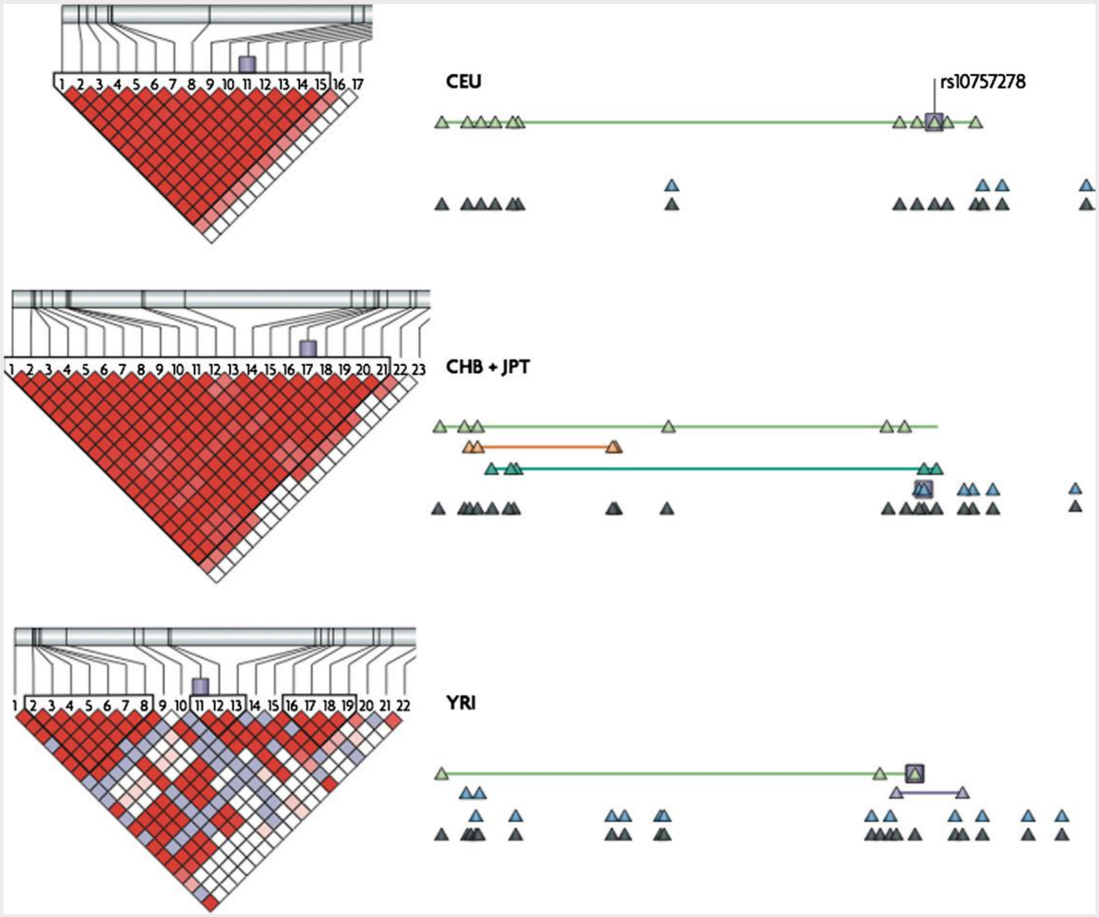
12.3.2 Linkage Studies

The purpose of linkage analysis is to search for alleles or chromosomal segments shared by affected relatives more so than expected by random Mendelian segregation. These segments are passed entirely from the parents to the offspring without recombination at meiosis. At gamete formation, the gamete receives a mixture of the two homologous chromosomes because of crossover events. Crossovers can split alleles that lie together on a common parental chromosome, resulting in alleles that originally came from different grandparents on the same chromosome. The number of crossover sites in the human genome is thought to follow a Poisson distribution (averaging around 35 crossover points), and their locations are generally random and independent of each other. Cosegregation should therefore be detectable for marker loci quite far away from the disease-causing variant. Because linkage operates over long genetic distances, a

BOX 12.3 Differences among Populations in the Linkage Disequilibrium of Common Variants in the Human Genome

The LD structure of SNPs in a 13 kb interval of chromosome 9p21 is shown in the accompanying figure for the three HapMap populations: CEU (European ancestry), JPT + CHB (Asian ancestry), and YRI (African ancestry). Two commonly used definitions of LD, D′ and r^2, capture different aspects of nonrandom association. On the left of the figure, the LD structures of the interval are shown quantified using D′. SNPs ascertained in phase II of the HapMap Project with a minor allele frequency (MAF) ≥5% are shown in their respective map positions. There are population differences in the numbers of SNPs that meet the ≥5% criterion. The pairwise correlations of SNPs—which are shown as vertical lines—appear as red and white boxes, with red indicating high correlation ($D' \approx 1$) and white indicating no correlation ($D' \approx 0$) (other colors represent intermediate values). On the right, all SNPs appear in the bottom row as black triangles. Above this row, the SNPs are grouped into bins at an $r^2 > 0.8$ (using the ldSelect algorithm). SNPs that are efficiently tagged by each other ($r^2 > 0.8$) are in the same color and are connected by a line. Singleton bins that do not tag any other SNPs appear as individual blue triangles.

Using both r^2 and D′, it is clear that LD is less strong in the African population than in the Caucasian and Asian populations. This reflects the fact that some haplotype patterns across the genome were lost in population bottlenecks associated with human migration out of Africa. Using the D′ statistic, fewer SNPs are correlated, as indicated by fewer red boxes and more white boxes. Using the r^2 statistic, Africans show a greater number of singleton SNPs and the LD bins have fewer numbers of variants and span shorter lengths. The r^2statistic reveals greater differences in pairwise correlations between SNPs in all three populations because the allele frequencies vary substantially among them. For example, in Caucasians, SNP rs10757278 (purple box), which is associated with myocardial infarction, lies in a bin of eleven SNPs; in Africans, rs10757278 is in LD with three of these SNPs, and in Asians it is in a singleton LD bin. For rs10757278, the MAF is the same in Caucasians and Asians (50%) but considerably less in Africans (5%).



Source: *Frazer et al.* [13]. *Reprinted with permission from Macmillan Publishers Ltd.*

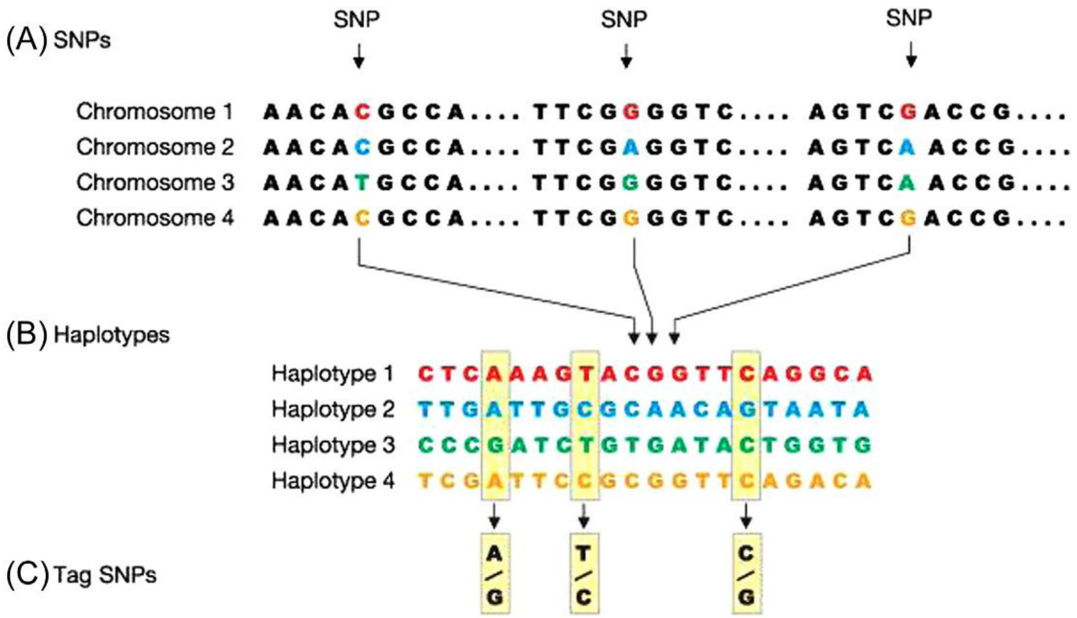
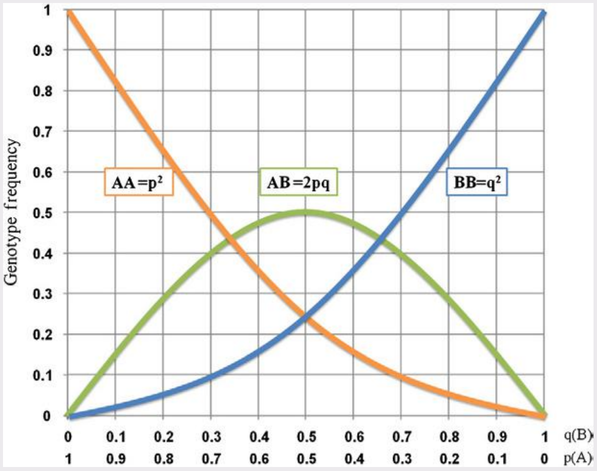

(A) SNPs
SNP
SNP
SNP
Chromosome 1 AACACGCCA.... TTCGGGGTC.... AGTCGACCG....
Chromosome 2 AACACGCCA.... TTCGAGGTC.... AGTCA ACCG....
Chromosome 3 AACATGCCA.... TTCGGGGTC.... AGTCA ACCG....
Chromosome 4 AACACGCCA.... TTCGGGGTC.... AGTCGACCG....
(B) Haplotypes
Haplotype 1 CTCAAAGTACGGTTCAGGCA
Haplotype 2 TTGATTGCGCAACAGTAATA
Haplotype 3 CCCGATCTGTGATACTGGTG
Haplotype 4 TCGATTCCGCGGTTCAGACA
(C) Tag SNPs
A/G
T/C
C/G

FIGURE 12.4 **Haplotype and tagged SNPs.** (A) A short stretch of DNA from four versions of the same chromosome region in different individuals. Most of the DNA sequence is identical in these chromosomes, but three bases are shown where variation occurs. Each SNP has two possible alleles; the first SNP in has the alleles cytosine and thymine. (B) A haplotype is made up of a particular combination of alleles at nearby SNPs. Shown here are the observed genotypes for 20 SNPs that extend across 6000 DNA bases. Only the variable bases are shown, which include the three SNPs that are shown in (A). For this region, most of the chromosomes in a population survey turn out to have haplotypes 1–4. (C) Genotyping just three tagged SNPs out of 20 is sufficient to uniquely identify these four haplotypes. For instance, if a particular chromosome has the pattern A-T-C at these three tagged SNPs, this pattern matches the pattern determined for haplotype 1. *Source: International HapMap Consortium* [34]. *Reprinted with permission from Macmillan Publishers Ltd.*

BOX 12.4 Hardy-Weinberg Equilibrium

The distribution of genotypes in a population in Hardy-Weinberg equilibrium can be graphically expressed as shown in the accompanying graph. The x-axis represents a range of possible relative frequencies of A or B alleles. The coordinates at each point on the three genotype lines show the expected proportion of each genotype at that particular starting frequency of A and B.



To check for HWE:

- Consider a single biallelic locus with two alleles A and B with known frequencies (allele A = 0.6; allele B = 0.4) that add up to 1.
- Possible genotypes: AA, AB and BB
- Assume that alleles A and B enter eggs and sperm in proportion to their frequency in the population (i.e., 0.6 and 0.4)
- Assume that the sperm and eggs meet at random (one large gene pool).
- Calculate the genotype frequencies as follows:
  - The probability of producing an individual with an AA genotype is the probability that an egg with an A allele is fertilized by a sperm with an A allele, which is 0.6 × 0.6 or 0.36 (the probability that the sperm contains A times the probability that the egg contains A).
  - Similarly, the frequency of individuals with the BB genotype can be calculated (0.4 × 04 = 0.16).
  - The frequency of individuals with the AB genotype is calculated by the probability that the sperm contains the A allele (0.6) times the probability that the egg contains the B allele (0.4), and the probability that the sperm contains the B allele (0.6) times the probability that the egg contains the A allele. Thus, the probability of AB individuals is (2 × 0.4 × 0.6 = 0.48).

Genotypes of the next generation can be given as shown in the accompanying table.

Allele	Allele Frequency	Genotype	Frequency	Counts for 1000
A (p)	0.6	AA	0.36	360
B (q)	0.4	AB	0.48	480
General formula of HW equation: $p^2+2pq+q^2=1$		BB	0.16	160
		Total	1	1000

The conclusions from HWE are follows:

- Allele frequencies in a population do not change from one generation to the next only as the result of assortment of alleles and zygote formation.
- If the allele frequencies in a gene pool with two alleles are given by p and q, the genotype frequencies is given by p2, 2pq, and q2.
- The HWE principle identifies the forces that can cause evolution.
- If a population is not in HWE, one or more of the five assumptions is being violated. Thus, HWE is based on five assumptions:
  Random selection: When individuals with certain genotypes survive better than others, allele frequencies may change from one generation to the next.
  No mutation: If new alleles are produced by mutation or if alleles mutate at different rates, allele frequencies may change from one generation to the next.
  No migration: Movement of individuals in or out of a population alters allele and genotype frequencies.
  No chance events: Luck plays no role in HWE. Eggs and sperm collide at the same frequencies as the actual frequencies of p and q. When this assumption is violated and by chance some individuals contribute more alleles than others to the next generation, allele frequencies may change. This mechanism of allele change is called genetic drift.
  Individuals select mates at random: If this assumption is violated, allele frequencies do change, but genotype frequencies may.

positional mapping approach based on linkage can cover the entire genome by using a relatively small number of highly polymorphic markers. Standard marker sets for whole-genome linkage scans, based on 200–800 microsatellite polymorphisms, which became available in the 1990s, enabled the successful mapping of hundreds of rare single-gene disorders.

Gene mapping makes use of recombination. The further apart two genes are, the higher the probability of an odd number of crossovers (odd numbers cause recombination), to a maximum of 50% (Figure 12.5). The recombination fraction (the proportion of meioses that result in recombination) is an indication of how far apart two genes are. It can be mathematically transformed into an expected number of crossover events. Distance along a chromosome is expressed in centimorgans. The relation between the length of DNA as measured in base pairs (bp) or centimorgans varies between men and women and from place to place in the genome, but a rule of thumb is that one centimorgan corresponds to about 1 billion bases.

Linkage analysis is carried out only in families with affected relatives and involves genotyping of several markers that spread over the entire genome. Markers that flank the disease gene or mutation tend to highly segregate with disease status in families. Identifying markers within such a segment that consistently accompany the disease may indicate the presence of susceptibility genetic factors near them. However, such factors are neither necessary nor sufficient for the disease to develop.

Linkage analysis can be parametric or nonparametric. Parametric analysis is so-called because it needs a precise assumption about the genetic model with defining multiple parameters. These include the mode of inheritance, gene frequencies, and penetrance. Parametric analysis is often performed on Mendelian traits, in which the genetic model can be easily specified. Complex common diseases, however, are much less tractable, making an advanced assumption impossible. For this reason, nonparametric or model-free linkage analysis is more applicable. In nonparametric analysis, DNA segments identical by descent among affected relatives are examined and analyzed mathematically in terms of the Mendelian probability of inheritance from the defined common ancestor. In the context of common disease, unaffected relatives offer negligible information for linkage analysis because of the low disease penetrance. Segments with excess IBD among affected relatives often harbor susceptibility factors.

12.3.2.1 Identity by Descent

DNA at the same locus on two homologous chromosomes is said to be identical by descent (IBD) if it originated from the same ancestral chromosome (Box 12.5). If two homologous chromosomes from different people are IBD at some locus, the people are related. If two homologous chromosomes from the same person are IBD at some locus, this person is inbred (i.e., has related parents). Two people, neither of whom is inbred, can share DNA IBD at a particular locus on either 0, 1, or 2 chromosomes. Identity by descent is useful for determining similarity between family members. IBD simply measures the amount of genetic material inherited by each offspring from each parent, and compares the two.

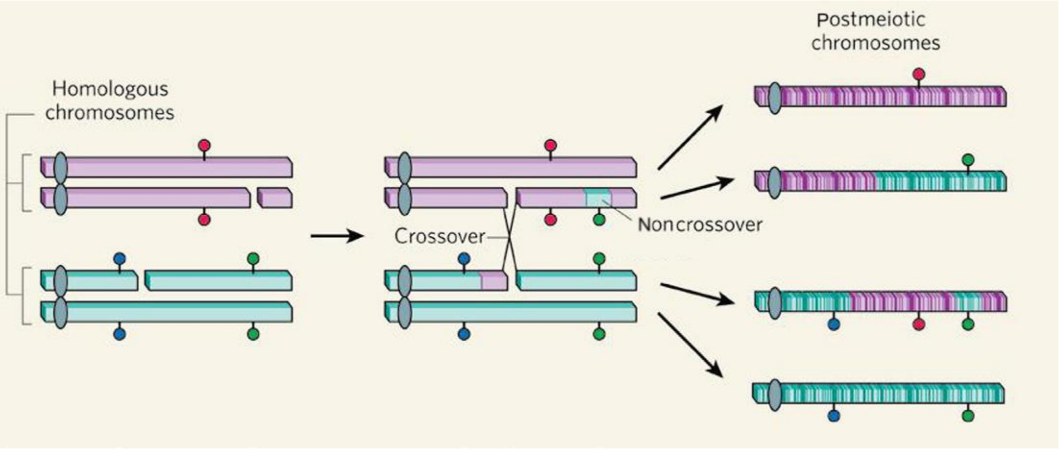


FIGURE 12.5 **Recombination and crossover during meiosis at gamete formation.** During meiosis, a DNA double-strand break is repaired by either crossover or noncrossover recombination, where both mechanisms erase LD. Different genetic markers (red, blue, and green lollipops) can be separated or combined by recombination into different post-meiotic chromosomes. For instance, the association between markers (blue and green) is disrupted by recombination in the second post-meiotic chromosomes because the green marker is no longer associated with the blue marker, whereas the association is conserved in the fourth post-meiotic chromosomes. *Source: From Lichten [35]. Reprinted with permission from Macmillan Publishers Ltd.*

Affected sib pair (ASP) study is one way to conduct a nonparametric linkage analysis. This supposes that if a given marker is co-segregating with a disease-predisposing allele then affected siblings of affected persons are more likely to receive the same allele identical by descent at a closely linked marker locus than if the marker locus was segregating independently (i.e., unlinked) of the disease-predisposing allele. Under the null hypothesis of no linkage, any pair of sibs are expected to share 0, 1, or 2 parental haplotypes with frequencies of 0.25, 0.5, and 0.25, respectively. In other words, according to Mendelian segregation rules, the probability that a pair of siblings share both marker alleles, one marker allele, or no marker alleles, IBD is 0.25, 0.5, and 0.25, respectively (see the accompanying table). However, if the marker is linked to a disease locus, the overall amount of sharing between affected sibs increases. For Mendelian diseases, linkage analysis has successfully identified the major genes and culpable gene variants. For instance, cystic fibrosis is found to be caused by multiple mutations in the CFTR genes. For common complex diseases, however, it has been much less successful. While several factors might explain this failure, the low heritability of most complex diseases and the low penetrance of causal variants are among the most likely reasons [19].

12.4 ASSOCIATION STUDIES

The goal of population association studies is to identify patterns of polymorphisms that vary systematically between individuals with different disease states and can therefore represent the effects of risk-enhancing or protective alleles [18]. This implies that traits are still linked to the surrounding genetic sequence of the original evolutionary ancestor through linkage disequilibrium, or that they are found more often in a given haplotype than outside of it. Association mapping is based on the idea that genetic variants underlying complex traits occur with a relatively high frequency (>1%), have undergone little or no selection in earlier populations, and are likely to date back >100,000 years (the common disease/common variant hypothesis). Association analysis potentially has far greater power than linkage analysis for detecting variants with a modest effect on disease risk, provided that the genetic marker is close enough to exhibit strong linkage disequilibrium (LD) with the functional variant (See Figure 12.6).

A direct association study tests the association between a known functional variant and disease. Indirect association studies are more commonly performed, which tests the association between the disease and a marker locus that lies close to the disease locus and is in LD with it. They may be enhanced by examining multiple markers simultaneously using haplotypes. A haplotype is a linear arrangement of closely linked alleles on the same chromosome that is inherited as a unit. An individual's genotypes at multiple, tightly linked SNPs resolve into two haplotypes, each containing alleles from one parent.

12.4.1 Association Study Designs

12.4.1.1 Candidate-Gene Studies

Candidate-gene studies depend on a set of markers based on *a priori* hypothesis about the role of a selected gene, or a group of pathway-related genes, on a phenotype. The single-gene approach offers the advantage that highly relevant genes can be prioritized and tested first. However, it places a

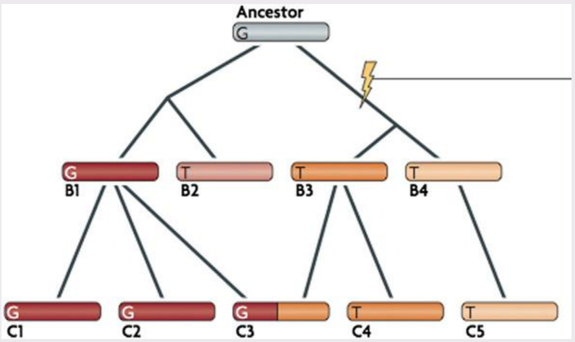
IBD Sharing among Different Relative Pairs

IBD Sharing at Single Locus	Parents	Parent-child	Full Siblings	Grandparent-grandchildren	Uncle-niece	Half Siblings	Identical Twins
Expected probability of 2 alleles, shared IBD	0	0	0.25	0	0	0	1
Expected probability of 1 allele, shared IBD	0	1	0.5	0.5	0.5	0.5	0
Expected probability of 0 alleles, shared IBD	1	0	0.25	0.5	0.5	0.5	0
Proportion of alleles with shared IBD	Exactly 0	Exactly 0.5	On average 0.5	On average 0.25	On average 0.25	On average 0.125	Exactly 1

Note: Assumption: unrelated parents.

BOX 12.5 IBD, IBS, and Coalescence

Here we illustrate the concepts of identity by descent (IBD), identity by state (IBS), and coalescence at a single nucleotide and at a chromosome segment. The accompanying figure depicts an ancestral allele at a locus representing the point of coalescence for alleles in the current population (C1–C5). At the point of coalescence (the most recent common ancestor), this locus carries a copy of a G allele that is subject to a mutation event (G → T; the lightning symbol), leading to a G/T polymorphism. IBD at the polymorphic locus among individuals (C1–C5) can be defined with respect to a base population (B1–B4) in which individuals are assumed to be unrelated (different color chromosome segments in the figure). Then the G alleles in C1, C2, and C3 are IBD to each other because all three descend from the G allele in B1. The T alleles in C4 and C5 are IBS but not IBD because they descend from different alleles in the base population.



Ancestral G allele at a given locus

Mutation event changing the G allele to a T

Base population used for the estimation of IBD coefficients

Current, observed, population of chromosome segments in five individuals

The whole chromosome segments C1 and C2 are IBD because they descend from a common ancestor (B1) without recombination, but chromosome segment C3 is not IBD to C1 and C2.

Source: *Powell et al.* [36]. *Reprinted with permission from Macmillan Publishers Ltd.*

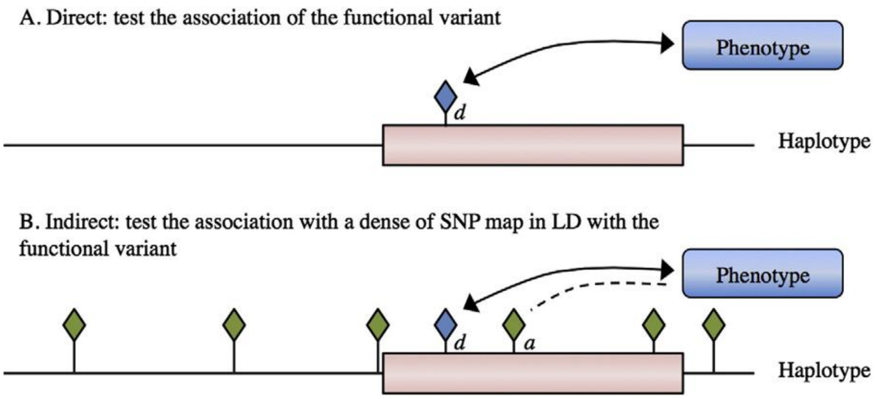

A. Direct: test the association of the functional variant
Phenotype
d
Haplotype
B. Indirect: test the association with a dense of SNP map in LD with the functional variant
Phenotype
d
a
Haplotype

FIGURE 12.6 **Direct and indirect association.** LD is the key factor in an association study because it induces correlation in short regions of the genome and underlines the susceptibility factors that have been inherited from ancient common ancestors. The d allele is associated with the disease. The a allele, on the other hand, is always associated with the d allele. Thus, an association with the disease can be directly found by genotyping the disease allele d (A), or indirectly found by genotyping the a allele that is in LD with the disease allele (B). Finding the direct association always results in higher association power, but indirect association should result in a significant association when the r^2 of alleles a and d is high and an adequate sample size has been used.

greater burden on the researcher to choose good candidates. The pathway-gene approach involves several candidate genes that together carry out related functions. Examples of pathways are genes related to drug metabolism (pharmacokinetics) or drug response (pharmacodynamics). A primary advantage of the candidate pathway-gene strategy is the ability to identify the effects of an aggregate of genes on a phenotype in circumstances where individual genes have small effects on the downstream phenotype. However, its success is heavily dependent on assumptions underlying the selection of genes to be studied. The risk is that unexpected genes that play an important role in drug pharmacokinetics and pharmacodynamics may remain undiscovered.

In general, experience with candidate-gene studies across all traits has been disappointing. In an extensive review, only 6 out of 166 putative associations were consistently replicated [9,10,20]. The failure of candidate-gene studies to identify the genetic basis of common traits suggests that this approach is subject to several limitations. First, the choice of candidate genes may be inappropriate. Second, the causative genes might be either upstream of the points of action or in the downstream signaling pathways of the selected candidates. Third, the SNPs selected may provide incomplete coverage of all variants in the genes under study. Fourth, most studies are underpowered and encounter problems with population stratification and phenotypic and locus heterogeneity. Finally, candidate-gene studies rely on prior hypotheses about disease mechanisms, which precludes discovery of genetic variants in previously unknown pathways.

The most comprehensive analysis of candidate genes is carried out by resequencing the entire gene in patients and controls and searching for a variant or set of variants that is enriched or depleted in disease cases. Such studies were previously laborious and expensive, and have been largely limited to one or a few candidate genes with sequencing restricted to the coding regions. However, with the advent of next-generation sequencing, whole exome sequencing in which all the coding regions of the genome is sequenced and variants tested for association is becoming more common—though this is technically not a candidate-gene study.

12.4.1.2 *Genome-Wide Association Studies*

Genome-wide association studies (GWAS) offer a hypothesis-free approach that systematically tests hundreds of thousands or more variants in the genome without prior knowledge of the location of the causal variants (Figure 12.7) [19,21]. GWAS are ideal for testing common variants with small effect sizes (Figure 12.8) and thus large sample sizes are required to generate sufficient statistical power. The high number of false-positive results they generate is addressed by stringent multiple testing correction and seeking evidence from multiple replication and validation studies of the top signals. There have been a series of successful GWAS studies of pharmacogenomic traits; these are summarized in Chapter 1.

12.4.2 Population-Based Study Designs

The most popular population-based study design for binary traits (whether focusing on candidate genes, regions, or the entire genome) is the case control study. In a case control study, a set of cases is identified and their genotyping information compared to a set of suitable controls in order to find genetic variants associated with the trait under study. Cases and controls are usually sampled independently from the same ethnic source population. An alternative approach involves sampling from a cohort of individuals that is being followed up prospectively (i.e., nested case control or case cohort study designs), such as provided by large-scale biobanking initiatives.

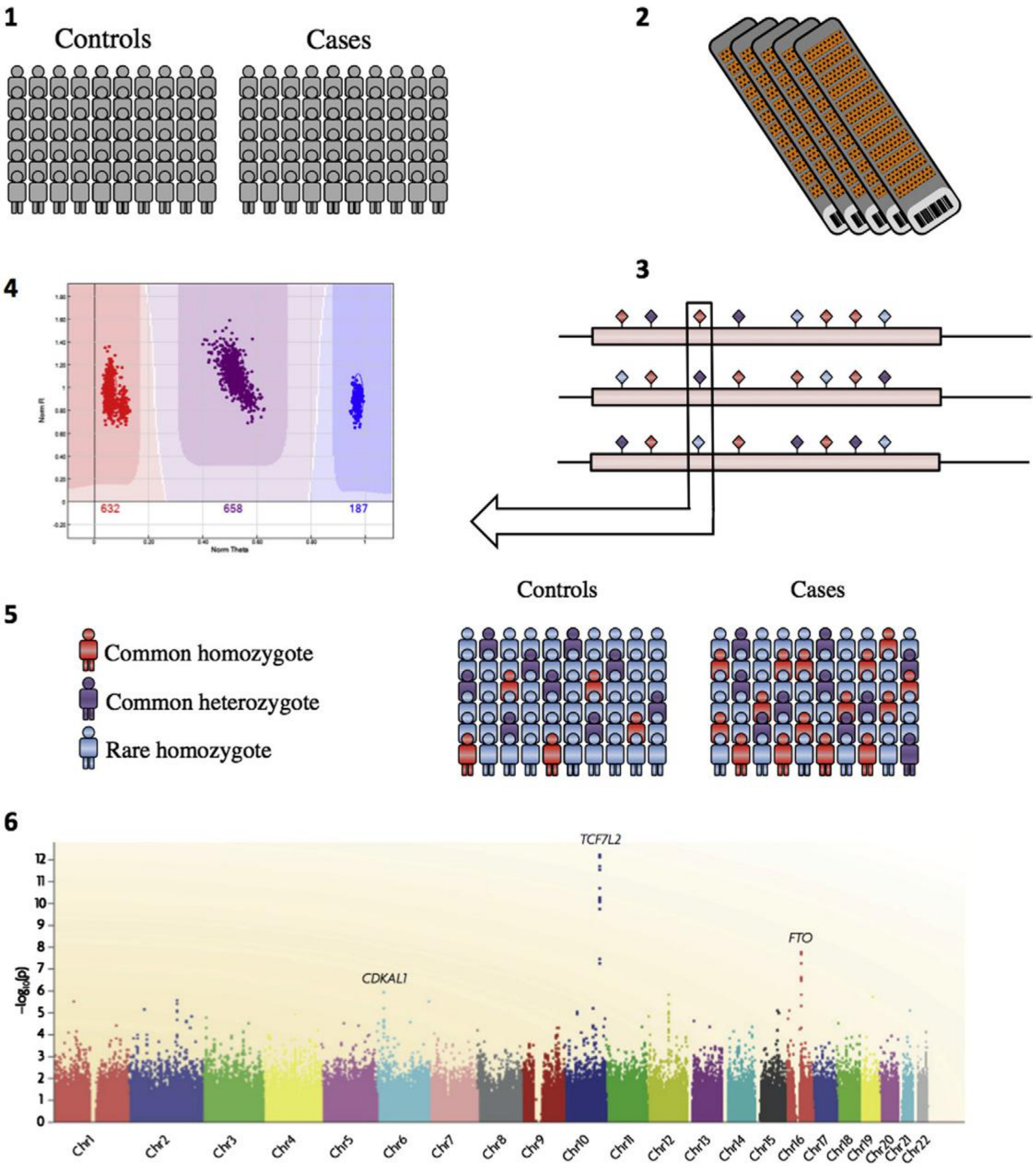

1
Controls
Cases
2
3
4
Norm R
Norm Theta
632
658
187
5
Common homozygote
Common heterozygote
Rare homozygote
Controls
Cases
6
−log10(p)
0
1
2
3
4
5
6
7
8
9
10
11
12
TCF7L2
FTO
CDKAL1
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Chr14
Chr15
Chr16
Chr17
Chr18
Chr19
Chr20
Chr21
Chr22

FIGURE 12.7 **Principles of GWA studies.** (1) Typically GWAS uses a case-control design where the study population is divided into two distinct groups on the basis of phenotype. (2) Samples are genotyped by the GWAS chip, which contains hundred of thousands of markers across the genome. (3) A representative DNA segment from three subjects contains several markers that are typed by the GWAS chip. (4) A cluster plot for signal intensity provides diagnostics for each marker across all samples; the raw data from the genotyping platform is plotted along two axes (one for each allele) to define clusters of data corresponding to the genotype groups. (5) The frequencies of each genotype are compared between cases and controls. (6) The results of the GWAS are represented as genome-wide Manhattan plots, which show the study findings with respect to their genomic position (each color represents a chromosome). The statistical significance threshold differs between studies, but in general signals with $p < 10^{-8}$ are considered significant. *Source: Manhattan plots adapted from McCarthy* [21] *with permission from Macmillan Publishers Ltd.*

For quantitative traits, individuals are either sampled from the two tails of the distribution and their genotyping information is compared similar to a case control study, or they are sampled randomly across the spectrum of trait values in a population; association is investigated by analyzing mean trait values between genotypes categories

12.4.2.1 Tests of Association

For association studies, the genotype effect can be modeled based on mendelian patterns or multiplicative/genotypic models (See Table 12.4). Test of association for continuous traits can be done by linear regression and analysis of variance (ANOVA) statistical methods. Both tests require a

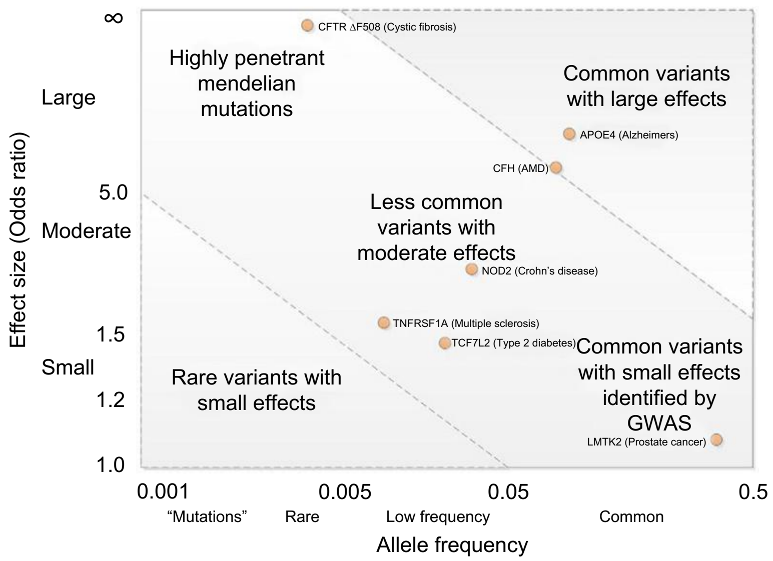


FIGURE 12.8 **Spectrum of allele frequency and effect size in genetic disease.** Rare variants with very large effect size (*upper left*) are the main cause of Mendelian diseases. The common complex diseases are associated with common SNPs with small effect size, which are targeted by GWAS (*lower right*). The bulk of discovered genetic associations lie on the diagonal denoted by the dashed lines. *Source: From Bush and Moore* [17].

TABLE 12.4 Linear Predictors for Genotype Variables Selected Standard Disease Models

Genotype	Model			
	Multiplicative	*Genotypic*	*Recessive*	*Dominant*
a/a	b_0	b_0	b_0	b_0
a/b	b_0+b_1	$b_0+b_1+b_2$	b_0	b_0+b_1
b/b	b_0+2b_1	b_0+2b_1	b_0+b_1	b_0+b_1
Interpretation				
Multiplicatives	b_1 provides an estimate of the log odds ratio for disease risk associated with each additional A allele (also called the haplotype relative risk). If b_1 is significant, there is a multiplicative contribution to disease risk in that the odds ratio for disease risk increases multiplicatively for every additional A allele.			
Genotypic	b_1 and b_2 provide estimates of the log odds ratio for disease risk in individuals with genotypes a/A and A/A, respectively, relative to an individual with genotype a/a. A likelihood ratio test of whether both b_1 and b_2 are significant is equivalent to the conventional 2 d.f. test for association in a 2 × 3 contingency table.			
Recessive	b_1 provides an estimate of the log odds ratio for disease risk in an individual with at least 1 A allele (genotype A/A or a/A) compared with an individual with no A alleles (genotype a/a). A test of whether b_1 is significant corresponds to a 1 d.f. test for association in a 2 × 2 contingency table of disease outcome by genotype classified as A/A or not.			
Dominant	b_1 provides an estimate of the log odds ratio for disease risk in an individual with two A alleles (genotype A/A) compared with an individual with one or two A alleles (genotype a/A or a/a). A test of whether b_1 is significant corresponds to a 1 d.f. test for association in a 2 × 2 contingency table of disease outcome by genotype classified as a/a or not.			

Source: Clarke et al. [22]. Reprinted with permission from Macmillan Publishers Ltd.

normal distribution of the trait. The linear regression analysis assumes a linear relationship between mean value of the trait and genotype with one degree of freedom. On the other hand, ANOVA is performed under the null hypothesis of no association with a general alternative and is analogous to the Pearson two degree-of-freedom test [18]. The genotype–phenotype relationship of quantitative traits can be expressed in a linear dose–response model in which the trait mean of the heterozygote lies midway between the two homozygotes. This model assumes that dominance is absent, and the strength of the genotype–phenotype relationship is expressed using a single parameter with the statistical test of association having one degree of freedom.

Tests of association are usually performed separately for each individual SNP. A biallelic SNP genotype and phenotype are represented in a 2×3 contingency table and a Pearson test (2 df) or a Fisher exact test to test the null hypothesis of no association between the rows and columns. Here, no order or trend of the genotype-phenotype relationship is assumed because each genotype is assumed to associate independently with the trait. This method allows testing of alternative models of association. If the genotype risks are additive, the Pearson test or a Fisher exact test will not be as powerful as tests, such as an allelic test, that are tailored to this scenario. An allelic test is carried out by counting alleles rather than genotypes so that each individual contributes twice to a 2×2 table and then applying a Pearson 1-df test. The allelic test of association is performed under the assumption of HWE in cases and controls combined and does not lead to interpretable risk estimates.

The Cochran-Armitage trend test is similar to the allelic test, but it is more conservative and does not rely on an assumption of HWE. Rather, it tests the hypothesis of zero slope for a line that best fits the three genotypic risk estimates [18]. Adopting the Armitage test implies sacrificing power if the genotypic risks are far from additive, in order to obtain better power for near-additive risks. A different approach is to adopt the Armitage test when the minor-allele frequency is low and the Fisher test when the counts for all three genotypes are high enough for it to have good power for nonadditive models [18].

Relative risk (RR) is a commonly used parameter to express the strength of association and compare disease penetrance among individuals with different genotypes. RR is only applicable for cohort studies where two groups (i.e., exposed and unexposed) can be followed to determine who develops the disease. For case-control studies, odd ratio (OR) is used instead of RR to measure the strength of the association, as it is not possible to make direct estimates of disease penetrance. The value of the OR of a disease is the probability that the disease is present compared with the probability that it is absent in exposed versus unexposed groups.

Although the odds of disease are not directly measurable because of sampling selection, conveniently, it is mathematically equivalent to the exposure OR (i.e., the odds of exposure in cases versus controls), which is easily calculated from exposure frequencies [22]. The allelic OR explains the association between disease and allele by comparing the odds of disease with the A allele to the odds of disease with the B allele. The genotypic ORs describe the association between disease and specific genotype by comparing the odds of disease with one genotype to the odds of disease with another genotype. Typically, two genotypic ORs can be calculated by comparing the two homozygous genotypes (i.e., AA to BB), and comparing the heterozygous genotype to the rare variant homozygous (i.e., AB to BB). The difference between RRs and ORs is very small for a disease with small penetrance. Furthermore, a multivariate statistical method, such as logistic regression, can be performed so that OR includes more SNPs, risk factors, and clinical variables.

12.4.2.2 *Study Power and Coverage*

The power of a genetic association study is its ability to find an association with significant statistical support, in which true findings can be distinguished from chance findings. GWAS power depends on sample size, allele frequency, effect size, and genotype platform coverage [23]. The theoretical expected effect sizes for common variants are for genotypic $RR < 2.0$ (common variants identified by GWA studies have genotypic RR between 1.1 and 1.5). The sample sizes required are further inflated by the large number of hypotheses that are tested in GWA studies, as p-values have to be corrected for multiple-hypothesis testing. Most of the GWA studies so far developed have analyzed variants with $MAF > 5\%$; to detect lower-frequency SNPs even larger sample sizes are required.

Genotyping platform coverage defines the extent of the genetic variation in the human genome that has been represented by markers on the platform [23]. Thus, r^2 represents an important parameter in association analysis because it is inversely proportional to the sample size required for detecting disease association [23]. For instance, if a typed marker is in LD ($r^2 = 0.5$) with untyped functional variants, the effective sample size to identify this association needs to be doubled. Thus, the power is reduced from 90% to less than 40% for a study of 6000 cases and 6000 controls and an MAF of 1%, to obtain a p-value on the order of 10^{-6} [23]. However, it is not reduced when using a marker with ($r^2 = 1$). Conventionally, a value of $r^2 > 0.8$ is considered sufficient for good coverage of untyped variants. This value allows genotyping of a lower number of markers with small losses in power. For a region with several common variants that are in LD, an iterative pairwise comparison can identify the set of tagged SNPs to be included on the platform (the lowest possible number of SNPs that cover the whole region at $r^2 > 0.8$). If the SNPs are in strong LD in this region, the number of tagged SNPs can be decreased by up to 70–80%. However, if they are in a region with low LD, almost every one might have to be genotyped to ensure comprehensive coverage of the region.

Several high-throughput genotyping platforms with different densities and selection methods are currently available: Affymetrix GeneChip, Illumina Infinium BeadChips, Perlegen, and Invader. Currently, marker densities range from 100,000 to more than 1 million per array. Markers are selected using one of four approaches: random selection, “tagging,” a combination of random selection and tagging, and selection of only SNPs with known functions. Arrays include a mixture of genetic variants that are polymorphic in certain populations with different ancestries. Thus, they can be used efficiently in different populations. The choice of genotyping platform is based on sample size, ancestral origin, and marker selection method. Generally, increasing marker densities up to 500,000 increases the study power if the sample size is adequate. However, 1 million–marker arrays are likely to have no great advantage over 500,000–marker arrays, except for African populations, which have lower LD.

A simple way to reduce genotyping costs while retaining power is to perform the study in two or three stages: first, genotype a large number of SNPs in a fraction of samples and then the use of a liberal p-value threshold to find a subset of SNPs with putative associations. Second, genotype the selected subset of SNPs from the first stage in the full sample or a larger one.

12.4.3 Correction for Multiple Testing

The power of association studies is the probability of their rejecting the null hypothesis of no association when the marker is truly associated with phenotype. A type I error is the probability of the null hypothesis being rejected when it is true—that is, a false-positive association between a phenotype and a marker is shown. For a single test, the conventional statistical significance level (α) is 0.05, which means that the probability of rejection of a truly positive null hypothesis is 5%. However, genetic association studies tend to study multiple markers, phenotypes, or the results of multiple genetic models or covariate adjustments [16,18,23]. Consequently, the cumulative probability of making type I errors increases over the entire GWA analysis because of multiple testing issues.

The family-wise error rate (FWER) is the probability of making one or more type I errors in all or part of the hypothesis when performing multiple hypothesis tests. Reducing FWERs decreases the probability of type I errors at the expense of reducing the power to detect truly positive association. In the study design stage, the number of performed statistical comparisons should be considered in order to set a suitable FWER. This means that the association significance level for each marker is corrected for multiple testing to maintain the overall FWER [24]. A correction for multiple testing can be performed using Bonferroni's adjustment, in which the total number of performed test (n) is multiplied by the conventional significance level (α). Another adjustment is Sidak's $(1-(1-\alpha)^{1/n})$. However, both corrections assume that the tests are independent, which is not entirely true because of LD between markers; both corrections are therefore very conservative in the context of correlated tests.

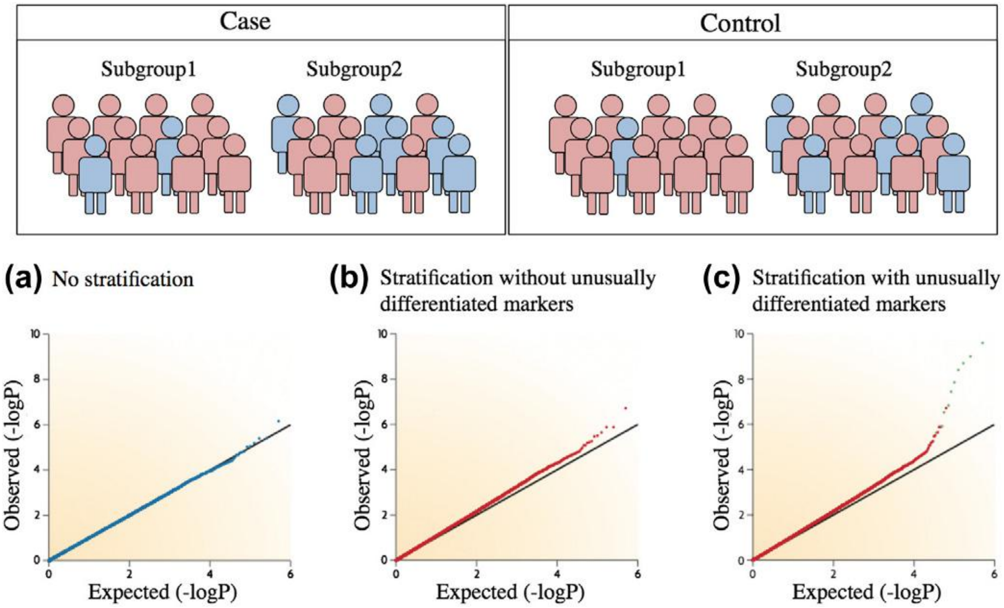
An alternative to FWER is permutation testing, which can be applied in typical association studies that involve huge numbers of tests. It is less conservative than FWER, as it considers the correlation between phenotypes and/or between genotypes. Permutation testing aims to make the null hypothesis correct by randomization; that is, the phenotype of each individual is randomly reassigned to another individual in the dataset and the association p-value is calculated. Hence, the genotype-phenotype correlation is broken except for correlation among genotypes and among phenotypes resulting from LD within an individual. Each reassignment step is repeated a predefined number of times (N) to produce an empirical distribution of p-values, which is then compared with the original p-value from the association test.

For GWAS, strong evidence of association is informally accepted to be below a standard genome-wide significance threshold of 5×10^{-8}. Moreover, plotting the observed p-values against the expected p-values in a quantile-quantile (QQ) plot is a useful for interpreting the GWAS results (Figure 12.9). The QQ plot is a graphical representation of the deviation of the observed p-values from the null hypothesis: the p-values for each SNP are sorted from largest to smallest and plotted against expected values from a theoretical χ^2-distribution. If the observed values correspond to the expected values, all points are on or near the diagonal between the x-axis and the y-axis. If some are clearly more significant than expected under the null hypothesis, points move towards the y-axis. An early separation of the expected from the observed means that many moderately significant p-values are more significant than expected under the null hypothesis. This result is rarely due to thousands of true positives; more often, it is due to population stratification: systematic differences in allele frequencies between subpopulations of the collection of individuals investigated, so that, by chance alone, a large number of p-values are smaller than expected.

12.4.3.1 Population Stratification and Admixture

Confounding by population origin or population stratification is an important and often quoted potential drawback of population-based genetic case control studies. Population stratification arises when (different proportions of) cases and controls are sampled from genetically different underlying populations, thus causing any associations found to be due to sampling differences rather than the disease of interest (Figure 12.9). Confounding occurs when such groups are unequally represented in the study and they differ in SNP allele frequencies and disease prevalence. Consequently, a biased or spurious association will be observed.

Several methods are established to detect and correct for population stratification. During the study design stage, population stratification can be minimized by matching controls to cases based on ethnicity or restricting inclusion to a specific ethnic group. This can be ascertained by



(a) No stratification **(b)** Stratification without unusually differentiated markers **(c)** Stratification with unusually differentiated markers

FIGURE 12.9 **Population stratification.** For a population that contains two subgroups, population stratification occurs when each subgroup has different SNP allele frequencies. The SNPs are thus erroneously associated with case or control status because of the population effect. Here, subgroup 2 has more blue alleles among cases but only because it is more frequent in this subgroup. P-P plots visualize stratification or other confounders. Shown are simulated P-P plots in three scenarios for genome-wide scans with no causal markers: (a) No stratification: p-values fit the expected distribution. (b) stratification without unusually differentiated markers: p-values exhibit modest genome-wide inflation. (c) Stratification with unusually differentiated markers: p-values exhibit modest genome-wide inflation and severe inflation at a small number of markers. *Source: from Price et al.* [26]. *Reprinted with permission from Macmillan Publishers Ltd.*

self-reported ethnicity. Moreover, if the disease prevalence differs between genders, matching based on sex may also reduce stratification.

The most common method used in GWAS is principal component analysis (PCA) using software such as EIGENSTRAT [25,26]. PCA is a multivariate statistical method to dissect and rank the correlation structures of different variables. The principal components (PCs) are calculated so that the first one has the maximum possible variance of its PC scores; this is followed by the second component and so on [25,26]. An alternative method is multidimensional scaling, which is implemented in PLINK but is computationally complex as it requires IBD estimation after merging the study samples with the reference populations. Both methods have to be performed on a large LD-pruned dataset that includes the study's sample and samples with known ancestry from a reference panel such as the HapMap2 or the HapMap3 phase III dataset. During the analysis stage, further correction for any residual subjects can be done by including the principal components as covariates in a logistic regression.

Another method is to use the genomic control to adjust for stratification in the absence of geographical information [26]. Genomic control assumes that the association test statistic is uniformly inflated by a constant factor (λ) and a higher median than expected. The magnitude of λ is estimated by comparing the observed median of a set of markers with their expected median under an assumption of no stratification. Population stratification is assumed to exist if λ has a value greater than 1, which can be corrected by dividing the association test statistic values by λ [13,22]. Genomic control has to be restricted to SNPs that are unrelated to disease status. For instance, in a candidate-gene study λ can be estimated only if the study includes an additional set of markers specifically selected to this purpose. For GWASs, however, the effect of putative SNPs on the estimation of λ is assumed to be negligible.

If geographical information about the study samples is known, then individuals can be classified into subgroups and the Cochran/Mantel-Haenszel test can be used to combine subgroups after separately analyzing them.

12.4.4 Imputation

Genotype imputation is a statistical technique often used to increase the power and resolution of genetic association studies. It works by using haplotype patterns in a reference panel to predict unobserved genotypes in a study dataset [27]. Imputation is typically used when combining data from genotyping platforms with different SNP contents. For instance, only 250,000 SNPs are shared by the Affymetrix SNP 6.0 and Illumina Human 1M arrays [28]. In combining the output of these two arrays to generate a new dataset with higher coverage, imputation reduces the batch effects and reduces direct genotyping costs. Several imputation methods are available [27]. A popular one uses a variant of the expectation maximization algorithm—an iterative frequentist method of maximizing the likelihood of the imputed variant. Basically, random variants are selected to fill the missing genotype, and the likelihood of observing this selection is examined. Then the imputed values are sequentially corrected over 10–100 iterations until the estimated error is minimized. The algorithm has to be run many times with different initial states to provide certainty, as it can get caught in a local maximum (providing a false impression of the best solution). An alternative is to use fitted Bayesian models using Markov chain Monte Carlo (MCMC) algorithms, which are iterative and explore the entire model space, not just the maximum.

12.4.5 Validation and Replication

Independent replication represents the sine qua non for reliably claiming a true genetic association [29]. A "true replication" indicates the evaluation of statistically significant association signals in additional independent study samples. Such replication efforts aim at reproducing the association signal identified in the initial study in nonoverlapping study samples, and will typically involve the same genetic variant, or a variant with very similar or identical statistical properties (a "proxy") as defined by a high correlation coefficient (r^2).

"Technical validation," conversely, indicates the reanalysis of all, or a subset of, genetic variants investigated in the association study and is carried out on the original study samples, often using a different genotyping platform. The aim of technical validation is to detect errors in genotyping or genotype imputation that could lead to differential genotype calling and spurious associations.

Evidence for "indirect replication" can be claimed when the replication study involves either a different allele at the same locus (e.g., different genetic variants or haplotypes in the same genomic region) or a phenotype trait that is closely correlated, but not identical to, the original study trait (e.g., an association with total cholesterol at a locus previously identified as associated with LDL cholesterol). While indirect replication can provide confirmatory evidence for the implication of a genetic locus in determining one or more traits, alone it does not provide sufficient evidence of replication of findings. However, often indirect replications can be important to gain additional insights into either the allelic architecture of a locus of interest, or of shared or unique genetic determinants of correlated biological traits or processes.

Failure to replicate the association signals on populations with different ancestries does not necessarily invalidate the initial findings. The differences in LD pattern across different populations can be used to narrow the region of interest for later genetic and, possibly, functional analysis [29]. An important factor to consider in a replication study is distinguishing between tests of the same SNPs as in the original study—those in strong LD with the reported SNP and others in the reported region—in order to search for additional variants association with the phenotype. For instance, a marker that is not in strong LD with a causal variant might be identified in a study; testing it in a different population might lead to false refutation of the association, but testing additional markers in the same region might reveal another association signal. Inclusion of new SNPs should be clearly rationalized by separating them from the other, previously tested SNPs. If the new SNPs are selected on the basis of LD patterns across populations with original SNPs, the different patterns should be empirically confirmed in the two populations—original and replication populations [29].

12.4.6 Meta-Analysis

Individual studies usually do not have enough power to detect variants with small effects due to small sample sizes. One way to increase power is by joint analysis (meta-analysis) of data from comparable GWA studies. Meta-analysis is used either for discovery of new variants or for replication of previous findings. Discovery analyses usually look across the whole genome, while replication analyses concentrate on a limited number of prespecified variants [30,31].

For the discovery meta-analysis, two designs are commonly used. The first is a consortium meta-analysis in which several partner research groups conduct GWA studies for the same phenotype. Each group follows the same criteria of quality control and primary analysis method. The groups can further collaborate to follow up any putative association signals. The second design involves combining the results of primary GWA studies with a publicly available GWAS data to increase study power. Typically, this design uses data from fewer primary studies that were initially intended to be independent but, because of the lack of power, were combined with other GWA studies. For this reason, difficulties can be encountered because of differences in quality control and method of analysis.

Replication meta-analysis studies aim to confirm previous findings as well as provide effect estimates that are free from self-reporting bias. A previously identified association's failure to be replicated may point to unsuspected interactions or, with some studies, methodological issues.

12.4.7 Fine Mapping and Resequencing

The ultimate goal of GWAS is to reveal functional variants underlying validated association signals and to characterize their effects. The confirmed signals that have passed all the discovery and replication stages indicate only statistical association. Furthermore, the typed markers are usually found in large LD blocks that may contain several genes, making it difficult to identify the causal gene. This scenario is even more complicated by the fact that many complex traits' susceptibility effects are mediated through remote regulatory elements, where the coding exon of the susceptibility gene might lie well beyond the interval of maximal association [21]. For these reasons, translating the confirmed association signals into an ordered list of candidate functional variants at a given locus is challenging. One way to identify functional variants is to perform fine mapping of the associated locus followed by deep resequencing of the association region. Candidate variants can then be prioritized by several bioinformatics and genomic methods to select a subset for further experimental study.

Fine-mapping approaches indicate a comprehensive exploration of variation around strong signals of association through additional genotyping and/or resequencing [32]. In fine mapping, additional genetic variants in regions of association are identified from sequence variant repositories (e.g., dbSNP, HapMap, 1000 Genomes project, www.1000genomes.org) or through targeted resequencing of association intervals (defined by recombination hotspot boundaries) in a subset of individuals to compile exhaustive catalogs of genetic variants within associated intervals. These variants are then statistically evaluated to distinguish between the true functional variants from other variants that are in LD with it. Selecting variants for fine mapping can sometimes be targeted more toward variants with particular aspects such as those with particularly strong biological credentials or those previously examined by expression studies [21]. This can facilitate the identification of true functional variants and confirm the implicated molecular mechanisms. In the absence of such clues, targeted resequencing is needed to examine the locus. Resequencing is also useful in identifying rare variants that are beyond the reach of typical GWA studies.

Targeted resequencing allows a comprehensive inventory of genetic variants within the targeted regions, leading to their systematic fine mapping. The objectives of resequencing have to be explicitly stated in order to formulate a study plan, which has to consider several factors: defining the ratio of reference samples, such as individuals from the HapMap panel, to cases, and determining if cases will be selected on the basis of carrying the known susceptible variant or haplotype; defining the extent of resequencing to be undertaken that will have an impact on the allele spectrum recovered; and, finally, determining if the resequencing will search for rare variants while acknowledging the difficulties in obtaining robust statistical evidence for rare variants in unannotated sequences [21].

12.4.8 Reporting of GWA Studies

Any report of GWA study results should include a detailed description of the study design, genotyping procedure and platforms, quality control procedures and thresholds, statistical methods and calculations, and results. Examples can be found in published papers or even as online supplements. A checklist containing key information that should be included in any GWA study report was recently published [29,33]. A checklist is necessary because the number of published GWA studies is steadily increasing, with wide variability in reporting and in the quality of their empirical findings, which may obscure their strengths and weaknesses. This in turn will have an impact on the scientific and clinical relevance of their results, in particular within the framework of systemic reviews and meta-analysis [29,33].

12.5 CONCLUSIONS

This chapter has provided an overview of the principles of determining the genetic basis of a phenotype. While the principles remain the same, technological advances during the last decade have revolutionized DNA sequencing. Next-generation sequencing generates huge amounts of data, which poses a challenge both for data storage and analysis; consequently, it often necessitates the use of powerful computing facilities and efficient algorithms. The three common applications of next-generation sequencing are (1) variant detection to identify genetic differences between two samples; (2) RNA-seq to determine the expression level of genes, as well as discover novel transcripts; and (3) ChIP-seq, which is a method of screening protein-DNA interactions. The analytic methods for these technologies are rapidly evolving and describing them is beyond the scope of this book.

REFERENCES

[1] John B, Lewis KR. Chromosome variability and geographic distribution in insects. Science 1966;152(3723):711–21.
[2] Gershon ES, Goldin LR. Clinical methods in psychiatric genetics. I. Robustness of genetic marker investigative strategies. Acta Psychiatr Scand 1986;74(2):113–8.
[3] Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy—a genomewide study. N Engl J Med 2008;359(8):789–99.
[4] Maitland ML, Bakris GL, Black HR, Chen HX, Durand JB, Elliott WJ, et al. Initial assessment, surveillance, and management of blood pressure in patients receiving vascular endothelial growth factor signaling pathway inhibitors. J Natl Cancer Inst 2010;102(9):596–604.
[5] Pirmohamed M, Aithal GP, Behr E, Daly A, Roden D. The phenotype standardization project: improving pharmacogenetic studies of serious adverse drug reactions. Clin Pharmacol Ther 2011;89(6):784–5.
[6] Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. Nat Rev Genet 2013;14(2):139–49.
[7] Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. Nat Rev Genet 2008;9(4):255–66.
[8] Almasy L. The role of phenotype in gene discovery in the whole genome sequencing era. Hum Genet 2012;131(10):1533–40.
[9] Dolan ME, Newbold KG, Nagasubramanian R, Wu X, Ratain MJ, Cook EH, et al. Heritability and linkage analysis of sensitivity to cisplatin-induced cytotoxicity. Cancer Res 2004;64(12):4353–6.

[10] Watters JW, Kraja A, Meucci MA, Province MA, McLeod HL. Genome-wide discovery of loci influencing chemotherapy cytotoxicity. Proc Natl Acad Sci USA 2004;101(32):11809–14.

[11] Birkenfeld AL, Jordan J, Hofmann U, Busjahn A, Franke G, Krüger N, et al. Genetic influences on the pharmacokinetics of orally and intravenously administered digoxin as exhibited by monozygotic twins. Clin Pharmacol Ther 2009;86(6):605–8.

[12] Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 2010;42(7):565–9.

[13] Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. Nat Rev Genet 2009;10(4):241–51.

[14] Palmer LJ, Cardon LR. Shaking the tree: mapping complex disease genes with linkage disequilibrium. Lancet 2005;366(9492):1223–34.

[15] International HapMap Consortium. A haplotype map of the human genome. Nature 2005;437(7063):1299–320.

[16] Cordell HJ, Clayton DG. Genetic association studies. Lancet 2005;366(9491):1121–31.

[17] Bush WS, Moore JH. Chapter 11: genome-wide association studies. PLoS Comput Biol 2012;8(12):e1002822.

[18] Balding DJ. A tutorial on statistical methods for population association studies. Nat Rev Genet 2006;7(10):781–91.

[19] Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005;6(2):95–108.

[20] Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet Med 2002;4(2):45–61.

[21] Mccarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008; 9(5):356–69.

[22] Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. Nat Protoc 2011;6(2):121–33.

[23] Wang WYS, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet 2005;6(2):109–18.

[24] Zondervan KT, Cardon LR. Designing candidate gene and genome-wide case-control association studies. Nat Protoc 2007;2(10):2492–501.

[25] Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38(8):904–9.

[26] Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. Nat Rev Genet 2010;11(7):459–63.

[27] Howie B, Marchini J, Stephens M, Chakravarti A. Genotype imputation with thousands of genomes. G3: Genes, Genomes, Genetics 2011;1(6):457–70.

[28] de Bakker PIW, Ferreira MAR, Jia X, Neale BM, Raychaudhuri S, Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet 2008;17(R2): R122–8.

[29] Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, et al. Replicating genotype-phenotype associations. Nature 2007;447(7145):655–60.

[30] Thompson JR, Attia J, Minelli C. The meta-analysis of genome-wide association studies. Brief Bioinform 2011;12(3):259–69.

[31] Evangelou E, Ioannidis JP. Meta-analysis methods for genome-wide association studies and beyond. Nat Rev Genet 2013;14(6):379–89.

[32] Ioannidis JPA, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. Nat Rev Genet 2009; 10(5):318–29.

[33] Janssens ACJW, Ioannidis JPA, Bedrosian S, Boffetta P, Dolan SM, Dowling N, et al. Strengthening the reporting of genetic risk prediction studies (GRIPS): explanation and elaboration. Eur J Clin Invest 2011;41(9):1010–35.

[34] Gibbs RA, Belmont JW, Hardenbol P, Willis TD, Yu F, Yang H, et al. The international HapMap project. Nature 2003;426(6968):789–96.

[35] Lichten M. Genomics: thoroughly modern meiosis. Nature 2008;454(7203):421–2.

[36] Powell JE, Visscher PM, Goddard ME. Reconciling the analysis of IBD and IBS in complex trait studies. Nat Rev Genet 2010;11(11):800–5.

This page is intentionally left blank

Chapter 13

Bioinformatics: Concepts, Methods, and Data

Scott W. Robinson,[a] **Avid M. Afzal,**[b] **and David P. Leader**[c]

[a]*Institute of Cardiovascular and Medical Sciences, BHF Cardiovascular Research Centre, University of Glasgow, UK,* [b]*Unilever Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, UK,* [c]*School of Life Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK*

Chapter Outline

13.1 Overview 259
13.2 Analysis of Sequences of Nucleic Acids and Proteins 260
13.2.1 Analysis of Sequences for Small Motifs 261
13.2.2 Comparison and Alignment of Two Sequences 261
13.2.3 Dot Matrix Comparison of Two Sequences 263
13.2.4 Multiple Sequence Alignment 263
13.2.5 Database Searching 264
13.3 Molecular Evolution and Phylogenetics 264
13.3.1 Tree Terminology 265
13.3.2 Principles of Tree Inference 266
13.3.3 Methods of Tree Inference 267
13.4 Molecular Graphics and Protein Modeling 269
13.4.1 Protein Stability 269
13.4.2 Protein Data Bank File Format 271
13.4.3 Protein Modeling 272
13.4.4 Homology Modeling 272
13.4.5 Fold Recognition 275
13.4.6 Threading 275
13.4.7 *De Novo* Structure Prediction 276
13.5 Bioinformatics Databases 276
13.5.1 General Sequence Databases 276
13.5.2 Gene Expression 277
13.5.3 Noncoding Sequences and Epigenomics 278
13.5.4 Proteins, Metabolites, and Pathways 279
13.5.5 Pharmacogenomics 280
13.5.6 Annotation 280
13.5.7 Common File Formats 280
Glossary 284
References 284

13.1 OVERVIEW

Progress in pharmacogenomics has accelerated rapidly in the past decade, driven for the most part by technological developments in molecular medicine. The next decade will see pharmacogenomics driven by sequencing as well as by collection of, linkage between, and access to increasing amounts of molecular and phenotypic patient data. There is clear evidence, especially from oncology, that the development of molecular understanding of drug response—in parallel with a medical informatics system that allows observation of linkages between clinical presentation and biomarker status using large datasets to accelerate predictive biomarker discovery—is of vital importance to the success of a targeted therapeutic.

The recognition that drug target biology may be extremely complex is important for appropriate biomarker identification. The key bottleneck is data analysis in the molecular (bioinformatics) and clinical (medical informatics) realms. Bioinformatics is at the intersection of molecular biology, computing, and statistics, and its main challenges are data storage and data analysis. The most common type of data are DNA, RNA, and protein sequences. This chapter provides an introductory overview of the field of bioinformatics.

To compare them, sequences must first be aligned, which may involve determination of a scoring metric and

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00013-X
© 2014 Elsevier Inc. All rights reserved.

penalties for gaps. When working with a large amount of data, finding an optimal multiple-sequence alignment is described as NP-complete, so we require "heuristic" methods as alternatives to exact methods. Phylogenetics uses the similarities/differences between sequences to infer relationships, and helps us better study evolution. Protein modeling may also make use of multiple-sequence alignments to infer the structure of proteins with an unknown conformation. For example, a sequence of unknown structure may be mapped to one of known structure and adjusted in regions where the sequence is different and chemically unlikely to form the same structure. Sequence analysis, phylogenetics, and protein modeling are related in many respects and have long been regarded as bioinformatics fields [1].

Because of recent developments in technology, a multitude of high-throughput techniques are now in common use. These techniques gather "omics" datasets, often used to screen for biomarkers or identify pathways that appear to be affected by a disease state. New methods have been developed for their analysis (described in other chapters), specific to proteomics, metabolomics, and so on. A great deal of these newly gathered data have been added to existing databases, and new databases have been created to accommodate new types of data.

This chapter describes some key concepts and methods in areas traditionally thought of as bioinformatics and discusses several useful databases. It includes a brief background and a discussion of the difficulties with data storage and continual database growth. Finally, file types common to IT and bioinformatics are described and contrasted to familiarize the reader with the files they should expect to work with and to provide an understanding of the utility of an appropriate file structure.

13.2 ANALYSIS OF SEQUENCES OF NUCLEIC ACIDS AND PROTEINS

This section considers some of the common types of computational analysis one is likely to perform on a newly obtained nucleic acid sequence and on the protein sequence(s) resulting from its conceptual translation. Our discussion is restricted to different types of sequence comparisons, although identification of simple motifs (including codons), being regarded as the most minimal type of comparison, is included. Other forms of analysis, such as secondary and tertiary structure and physical properties, fall outside the scope of this section, as does the analysis of large genomic regions.

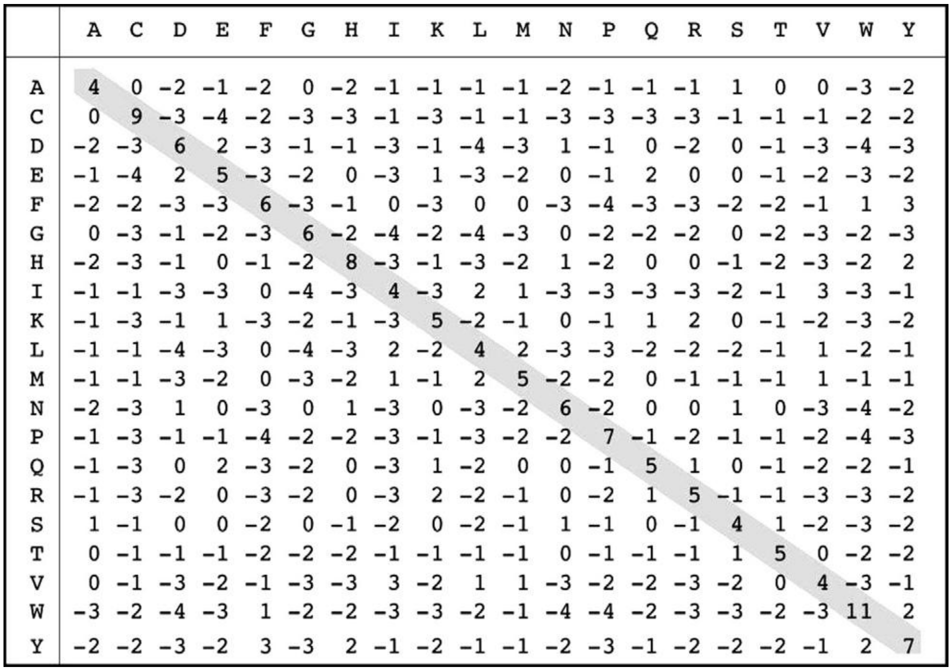
Before discussing sequence comparisons, it is pertinent to reflect on the metrics of comparison. Although this is trivial in the case of tasks requiring perfect matches (e.g., finding restriction endonuclease recognition sites in a DNA sequence), determining the best alignment of two (or more) similar but nonidentical sequences, say, involves quantifying similarities and differences.

In comparing two nucleic acid sequences, it might be reasonable to score 1 for a match and 0 for a mismatch between nucleotide bases, but such a simplistic approach is unsatisfactory for protein sequences. Experimental analysis shows that in the corresponding positions of closely related proteins certain amino acid changes are more common than others, which is understandable both in terms of the physicochemical requirements for amino acids in the three-dimensional structure of proteins (and in their function) and in terms of the different number of mutations that may be required to interconvert different amino acids. This latter consideration also implies that the scoring system one uses should take into account the time since the two sequences under comparison diverged from a common ancestor. Amino acid comparison matrices are available for different degrees of divergence, the BLOSSUM matrices [2] being in most common use today. Users should check that the default matrix used by a program is the most suitable one for the evolutionary distance between the sequences being compared (Figure 13.1).

In cases where a program presents a "best" sequence alignment, it is generally necessary to introduce gaps in it at positions where deletion or insertion mutations are presumed to have occurred. However, to prevent an unrealistic number of gaps, programs exact penalties for their introduction, so they are inserted only if the penalty is compensated for by the additional scoring matches it produces. Users should be aware that gap penalties may be varied, and that default values, although perhaps appropriate in most cases, may not be suitable for all situations (e.g., the ends of incomplete sequences).

In many cases there are a variety of programs, both commercial and free, that perform the comparisons to be discussed. Rather than trying to produce a list, we will refer primarily to programs of the open-source EMBOSS package [3], which can be either downloaded from http://emboss.sourceforge.net/ and run on the user's own computer (albeit primarily on the command line), or used on the web pages of EBI, the European Bioinformatics Institute (www.ebi.ac.uk/Tools/emboss/).

We will mention the general utility and limitations of programs in each category, but one aspect of the latter needs to be explained for the benefit of readers without a background in informatics. In writing software to solve a problem such as sequence alignment, the programmer tries to devise a logical procedure (algorithm) for achieving this and then implements it in computer code. The time it takes the coded program to run is related to the number of key operations, n, and generally in sequence comparison the more nucleotides or amino acids in a sequence, the more key operations have to be performed. What is important to realize is that the mathematical nature of the relationship between n and



	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
A	4	0	-2	-1	-2	0	-2	-1	-1	-1	-1	-2	-1	-1	-1	1	0	0	-3	-2
C	0	9	-3	-4	-2	-3	-3	-1	-3	-1	-1	-3	-3	-3	-3	-1	-1	-1	-2	-2
D	-2	-3	6	2	-3	-1	-1	-3	-1	-4	-3	1	-1	0	-2	0	-1	-3	-4	-3
E	-1	-4	2	5	-3	-2	0	-3	1	-3	-2	0	-1	2	0	0	-1	-2	-3	-2
F	-2	-2	-3	-3	6	-3	-1	0	-3	0	0	-3	-4	-3	-3	-2	-2	-1	1	3
G	0	-3	-1	-2	-3	6	-2	-4	-2	-4	-3	0	-2	-2	-2	0	-2	-3	-2	-3
H	-2	-3	-1	0	-1	-2	8	-3	-1	-3	-2	1	-2	0	0	-1	-2	-3	-2	2
I	-1	-1	-3	-3	0	-4	-3	4	-3	2	1	-3	-3	-3	-3	-2	-1	3	-3	-1
K	-1	-3	-1	1	-3	-2	-1	-3	5	-2	-1	0	-1	1	2	0	-1	-2	-3	-2
L	-1	-1	-4	-3	0	-4	-3	2	-2	4	2	-3	-3	-2	-2	-2	-1	1	-2	-1
M	-1	-1	-3	-2	0	-3	-2	1	-1	2	5	-2	-2	0	-1	-1	-1	1	-1	-1
N	-2	-3	1	0	-3	0	1	-3	0	-3	-2	6	-2	0	0	1	0	-3	-4	-2
P	-1	-3	-1	-1	-4	-2	-2	-3	-1	-3	-2	-2	7	-1	-2	-1	-1	-2	-4	-3
Q	-1	-3	0	2	-3	-2	0	-3	1	-2	0	0	-1	5	1	0	-1	-2	-2	-1
R	-1	-3	-2	0	-3	-2	0	-3	2	-2	-1	0	-2	1	5	-1	-1	-3	-3	-2
S	1	-1	0	0	-2	0	-1	-2	0	-2	-1	1	-1	0	-1	4	1	-2	-3	-2
T	0	-1	-1	-1	-2	-2	-2	-1	-1	-1	-1	0	-1	-1	-1	1	5	0	-2	-2
V	0	-1	-3	-2	-1	-3	-3	3	-2	1	1	-3	-2	-2	-3	-2	0	4	-3	-1
W	-3	-2	-4	-3	1	-2	-2	-3	-3	-2	-1	-4	-4	-2	-3	-3	-2	-3	11	2
Y	-2	-2	-3	-2	3	-3	2	-1	-2	-1	-1	-2	-3	-1	-2	-2	-2	-1	2	7

FIGURE 13.1 **BLOSSUM62 amino acid substitution matrix.** The central diagonal is shaded to emphasize how the score for a perfect match varies with amino acid. The "62" in the matrix name indicates that it was constructed from blocks with 62% identity. For evolutionary distant sequences, matrices constructed from blocks with lower identity, such as BLOSSUM45, are available.

the time a program takes to run varies depending on the algorithm: the time complexity (as it is called) can be on the order of logn, n, nlogn, n^2, n^3, 2^n, and so on.

The order of the time complexity is generally of much more importance than the length of the sequence, although the latter can sometimes be relevant. This means that there are situations in which there is a mathematically sound algorithm for solving a problem, but it would take an eternity to run. The threshold of practicality occurs at a time complexity of the order of n^2, and above this programs adopt algorithms employing what are called heuristics. This is an approach in which certain simplifying assumptions are made on the basis of what is usually the case (the "rule of thumb") to achieve an answer that is likely, but not certain, to be correct in a reasonable time. The user of sequence comparison programs employing this approach should be aware of this limitation, especially when interpreting unexpected results.

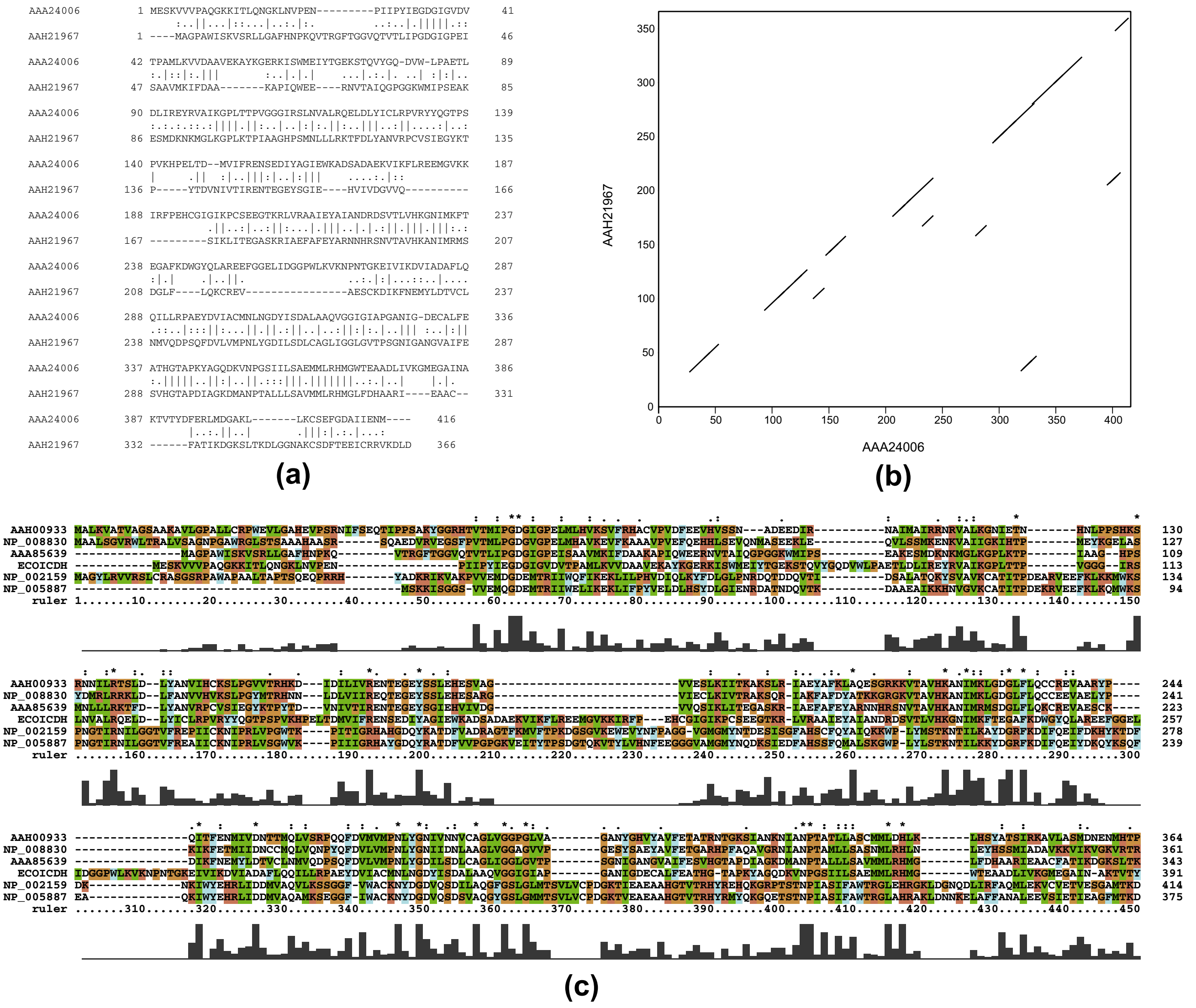
13.2.1 Analysis of Sequences for Small Motifs

The simplest type of sequence comparison is employed in the identification of small motifs such as restriction endonuclease recognition sites. The comparisons involved are exact, so no scoring system is involved, and the number of comparisons is of the order of the length of the sequence so that the program can run quickly. Even when a set of comparisons has to be made, in the case of searches for motifs, such as promoters with consensus sequences, time of execution is not a problem. A range of such programs is available in the EMBOSS package mentioned earlier.

A more sophisticated type of query involves searching a sequence for a pattern: a combination of motifs (either precise or consensus) with additional fuzzy features, such as that the distance between them falls within a range. This sort of pattern specification can be represented in a notation termed a "regular expression" in computing science, and processing methods for it exist in many computing languages, most notably Perl. For those without the requisite skills to write and use regular expressions, there are web facilities such as Prosite (http://prosite.expasy.org/scanprosite/) where one can enter a pattern in response to simple instructions, generating a regular expression that the underlying software uses to analyze the sequence of interest.

13.2.2 Comparison and Alignment of Two Sequences

It is often necessary to compare a sequence of interest (the query) with a second sequence that is believed to be related, in order to obtain the best detailed alignment of the two. This involves matching amino acid with amino acid (or base with base) and introducing gaps where appropriate, as illustrated in Figure 13.2(a). As this requires determination of the highest-scoring alignment from a huge number of



```
AAA24006           1 MESKVVVPAQGKKITLQNGKLNVPEN---------PIIPYIEGDGIGVDV     41
                         :..||...|::...|..:.|:.         ..:..|.|||||.::
AAH21967           1 ----MAGPAWISKVSRLLGAFHNPKQVTRGFTGGVQTVTLIPGDGIGPEI     46

AAA24006          42 TPAMLKVVDAAVEKAYKGERKISWMEIYTGEKSTQVYGQ-DVW-LPAETL     89
                     :.|::|:.|||       :..|.|.|    ...|.:.|. ..| :|:|..
AAH21967          47 SAAVMKIFDAA-------KAPIQWEE----RNVTAIQGPGGKWMIPSEAK     85

AAA24006          90 DLIREYRVAIKGPLTTPVGGGIRSLNVALRQELDLYICLRPVRYYQGTPS    139
                     :.:.:.:.::.:||||.||:..|..|:|:.||:..|||..:||....:|..:
AAH21967          86 ESMDKNKMGLKGPLKTPIAAGHPSMNLLLRKTFDLYANVRPCVSIEGYKT    135

AAA24006         140 PVKHPELTD--MVIFRENSEDIYAGIEWKADSADAEKVIKFLREEMGVKK    187
                     |     .||  :|..|||:|..|:|||    ....:.|::
AAH21967         136 P-----YTDVNIVTIRENTEGEYSGIE----HVIVDGVVQ----------    166

AAA24006         188 IRFPEHCGIGIKPCSEEGTKRLVRAAIEYAIANDRDSVTLVHKGNIMKFT    237
                              .||..:|..:||:...|.|||..|.|.:||.|||.|||:.:
AAH21967         167 ---------SIKLITEGASKRIAEFAFEYARNNHRSNVTAVHKANIMRMS    207

AAA24006         238 EGAFKDWGYQLAREEFGGELIDGGPWLKVKNPNTGKEIVIKDVIADAFLQ    287
                     :|.|    .|..||.                ..:.|:|...::..|....
AAH21967         208 DGLF----LQKCREV----------------AESCKDIKFNEMYLDTVCL    237

AAA24006         288 QILLRPAEYDVIACMNLNGDYISDALAAQVGGIGIAPGANIG-DECALFE    336
                     .:..|:::||:...||.||.:||..|..:||:|:.|..||| :..|:||
AAH21967         238 NMVQDPSQFDVLVMPNLYGDILSDLCAGLIGGLGVTPSGNIGANGVAIFE    287

AAA24006         337 ATHGTAPKYAGQDKVNPGSIILSAEMMLRHMGWTEAADLIVKGMEGAINA    386
                     :.|||||..||:|..||.:::|||.|||||||..:.|..|    |.|.
AAH21967         288 SVHGTAPDIAGKDMANPTALLLSAVMMLRHMGLFDHAARI----EAAC--    331

AAA24006         387 KTVTYDFERLMDGAKL-------LKCSEFGDAIIENM----    416
                           |..:.||..|       .|||:|.:.|...:
AAH21967         332 ------FATIKDGKSLTKDLGGNAKCSDFTEEICRRVKDLD    366
```

```
                                                                  :  : ** :   :   : . .  .         .:        .              :      .  :.:     *                *
 AAH00933 MALKVATVAGSAAKAVLGPALLCRPWEVLGAHEVPSRNIFSEQTIPPSAKYGGRHTVTMIPGDGIGPELMLHVKSVFRHACVPVDFEEVHVSSN---ADEEDIR----------NAIMAIRRNRVALKGNIETN-------HNLPPSHKS   130
NP_008830 MAALSGVRWLTRALVSAGNPGAWRGLSTSAAAHAASR------SQAEDVRVEGSFPVTMLPGDGVGPELMHAVKEVFKAAAVPVEFQEHHLSEVQNMASEEKLE----------QVLSSMKENKVAIIGKIHTP-------MEYKGELAS   127
 AAA85639 ---------------MAGPAWISKVSRLLGAFHNPKQ--------VTRGFTGGVQTVTLIPGDGIGPEISAAVMKIFDAAKAPIQWEERNVTAIQGPGGKWMIPS---------EAKESMDKNKMGLKGPLKTP-------IAAG--HPS   109
  ECOICDH -----------MESKVVVPAQGKKITLQNGKLNVPEN-----------------PIIPYIEGDGIGVDVTPAMLKVVDAAVEKAYKGERKISWMEIYTGEKSTQVYGQDVWLPAETLDLIREYRVAIKGPLTTP-------VGGG--IRS   113
NP_002159 MAGYLRVVRSLCRASGSRPAWAPAALTAPTSQEQPRRH-------YADKRIKVAKPVVEMDGDEMTRIIWQFIKEKLILPHVDIQLKYFDLGLPNRDQTDDQVTI---------DSALATQKYSVAVKCATITPDEARVEEFKLKKMWKS   134
NP_005887 ----------------------------------------------MSKKISGGS-VVEMQGDEMTRIIWELIKEKLIFPYVELDLHSYDLGIENRDATNDQVTK---------DAAEAIKKHNVGVKCATITPDEKRVEEFKLKQMWKS    94
    ruler 1.......10........20........30........40........50........60........70........80........90.......100.......110.......120.......130.......140.......150

           :  :*  :.  ::             .   .     :   *.   . * . :                                    :   .   :   :.. .:. *  .    :    * .*::  :* * :   ::
 AAH00933 RNNILRTSLD--LYANVIHCKSLPGVVTRHKD---IDILIVRENTEGEYSSLEHESVAG--------------------------VVESLKIITKAKSLR-IAEYAFKLAQESGRKKVTAVHKANIMKLGDGLFLQCCREVAARYP----   244
NP_008830 YDMRLRRKLD--LFANVVHVKSLPGYMTRHNN---LDLVIIREQTEGEYSSLEHESARG--------------------------VIECLKIVTRAKSQR-IAKFAFDYATKKGRGKVTAVHKANIMKLGDGLFLQCCEEVAELYP----   241
 AAA85639 MNLLLRKTFD--LYANVRPCVSIEGYKTPYTD---VNIVTIRENTEGEYSGIEHVIVDG--------------------------VVQSIKLITEGASKR-IAEFAFEYARNNHRSNVTAVHKANIMRMSDGLFLQKCREVAESCK----   223
  ECOICDH LNVALRQELD--LYICLRPVRYYQGTPSPVKHPELTDMVIFRENSEDIYAGIEWKADSADAEKVIKFLREEMGVKKIRFP---EHCGIGIKPCSEEGTKR-LVRAAIEYAIANDRDSVTLVHKGNIMKFTEGAFKDWGYQLAREEFGGEL   257
NP_002159 PNGTIRNILGGTVFREPIICKNIPRLVPGWTK----PITIGRHAHGDQYKATDFVADRAGTFKMVFTPKDGSGVKEWEVYNFPAGG-VGMGMYNTDESISGFAHSCFQYAIQKKWP-LYMSTKNTILKAYDGRFKDIFQEIFDKHYKTDF   278
NP_005887 PNGTIRNILGGTVFREAIICKNIPRLVSGWVK----PIIIGRHAYGDQYRATDFVVPGPGKVEITYTPSDGTQKVTYLVHNFEEGGGVAMGMYNQDKSIEDFAHSSFQMALSKGWP-LYLSTKNTILKKYDGRFKDIFQEIYDKQYKSQF   239
    ruler .......160.......170.......180.......190.......200.......210.......220.......230.......240.......250.......260.......270.......280.......290.......300

                          .*  :    *    : :     : : .  *  *:  .:  *   *. *: .       . .     .           .      .** : :::    * *             :  :       :.     .
 AAH00933 ----------------QITFENMIVDNTTMQLVSRPQQFDVMVMPNLYGNIVNNVCAGLVGGPGLVA-------GANYGHVYAVFETATRNTGKSIANKNIANPTATLLASCMMLDHLK-------LHSYATSIRKAVLASMDNENMHTP   364
NP_008830 ----------------KIKFETMIIDNCCMQLVQNPYQFDVLVMPNLYGNIIDNLAAGLVGGAGVVP-------GESYSAEYAVFETGARHPFAQAVGRNIANPTAMLLSASNMLRHLN-------LEYHSSMIADAVKKVIKVGKVRTR   361
 AAA85639 ----------------DIKFNEMYLDTVCLNMVQDPSQFDVLVMPNLYGDILSDLCAGLIGGLGVTP-------SGNIGANGVAIFESVHGTAPDIAGKDMANPTALLLSAVMMLRHMG-------LFDHAARIEAACFATIKDGKSLTK   343
  ECOICDH IDGGPWLKVKNPNTGKEIVIKDVIADAFLQQILLRPAEYDVIACMNLNGDYISDALAAQVGGIGIAP-------GANIGDECALFEATHG-TAPKYAGQDKVNPGSIILSAEMMLRHMG-------WTEAADLIVKGMEGAIN-AKTVTY   391
NP_002159 DK-------------NKIWYEHRLIDDMVAQVLKSSGGF-VWACKNYDGDVQSDILAQGFGSLGLMTSVLVCPDGKTIEAEAAHGTVTRHYREHQKGRPTSTNPIASIFAWTRGLEHRGKLDGNQDLIRFAQMLEKVCVETVESGAMTKD   414
NP_005887 EA-------------QKIWYEHRLIDDMVAQAMKSEGGF-IWACKNYDGDVQSDSVAQGYGSLGMMTSVLVCPDGKTVEAEAAHGTVTRHYRMYQKGQETSTNPIASIFAWTRGLAHRAKLDNNKELAFFANALEEVSIETIEAGFMTKD   375
    ruler .......310.......320.......330.......340.......350.......360.......370.......380.......390.......400.......410.......420.......430.......440.......450
```

FIGURE 13.2 **Different sequence comparisons illustrated for various examples of isocitrate dehydrogenase.** (a) Global alignment using the program Needle. The two amino acid sequences labeled "AAA2406" and "AAH21967" are of *Escherichia coli* ICDH and Human ICDH 3α, respectively. Identical amino acids are indicated by vertical bars, related amino acids by a double dot, and unrelated amino acids by a single dot. (b) Dot-matrix comparison using the program Dotmatcher. The sequences are as in (a). (c) Multiple-sequence alignment using the program ClustalX. Various human isoforms of ICDH are compared with ICDH from *E. coli.* Similar amino acids are highlighted by color; identical amino acids by an asterisk. The height of each bar corresponds to the number of identical amino acids at that particular position. (A final small section of the comparison is omitted for reasons of space.)

possibilities, at first sight it appears to fall into the category of problems that cannot be solved in a finite time. However, a sophisticated computational approach called dynamic programming allows the problem to be reduced to the generation of a two-dimensional (2D) matrix of pointers, the path through which represents the sequence alignment. This gives a time complexity on the order of n^2, which, although not ideal, allows sequences of up to 1000 or so amino acids to be aligned in a reasonable time.

The original application of dynamic programming to the complete alignment of sequences (global alignment) is due to Needleman and Wunsch [4], with a variant by Smith and Waterman [5] to allow only the most related regions to be aligned (local alignment). The EMBOSS implementation of these algorithms is entitled *Needle* or *Water*, respectively. A comparison matrix is used in the scoring system of these programs, as well as gap penalties of a little more complexity than previously indicated. Because insertion and deletion mutations often involve several amino acids, it is reasonable in determining gap penalties to distinguish between the creation of a new gap and its extension, the latter usually incurring a much smaller penalty than the former.

It is important to realize that it is not cheating to alter the penalties, or even to customize the comparison matrix. Default values are only the product of general biological experience and are not intrinsic to the algorithmic logic. If one has specific biological knowledge that can inform a particular sequence comparison, it is perverse not to use it (and, of course, report that one has done so). For example, if it is known that certain residues in a sequence should align because they represent, say, active site residues in an enzyme, it is perfectly reasonable to force their alignment. This can be done by modifying the comparison matrix to accommodate an unused symbol, "Z," assigning a high score to a perfect match with it, and replacing the residues to be aligned in the sequence with "Z."

One other feature of such alignments should be noted. It is possible for there to be more than one alignment of two sequences with the same score, although only one is presented in the output.

13.2.3 Dot Matrix Comparison of Two Sequences

Sometimes the object of comparing two sequences is not to obtain a base-by-base or amino acid-by-amino acid alignment, but to get an overall view of regions of similarity and difference. In such cases, dot-matrix comparison can be useful. The principle here is to place the two sequences on the axes of a 2D grid and draw imaginary lines horizontally and vertically back from each grid cell to the axes. If the two bases (or amino acids) at the points of intersection on the axes are identical, a dot is placed in the grid cell. For regions of high identity, the eye will see the dots as a continuous diagonal line, interrupted by insertions or deletions in one of the sequences (Figure 13.2(b)). The time complexity of this can readily be seen to be of the order of n^2.

In practice, a more sophisticated approach is adopted to determine whether a dot should be drawn at a particular position, so as to avoid a large amount of random noise in some cases (there is always a 1-in-4 chance of a match with nucleic acid sequences) or a signal too low to display similarities in other cases (e.g., diverged proteins). What programs such as Dotmatcher in the EMBOSS suite do is compare not a single base or amino acid but a short sequence window (longer for a nucleic acid sequence than for a protein sequence) and specify a threshold—a percentage of the maximum possible score—above which a dot should be drawn. In the case of proteins, a scoring matrix is used to determine the score. The importance of varying window and threshold values in this sort of program cannot be overemphasized—individual comparisons differ so much that default values frequently result in suboptimal signal-to-noise ratios.

It should be pointed out that, although dot matrix comparison is useful for revealing insertions and deletions (and for revealing intron positions in comparisons of genomic and cDNA sequences), gap penalties do not feature in the scoring system. All the program does is calculate whether or not to draw a dot—interpretation in terms of gaps and horizontal lines is done by the user when inspecting the resulting graph.

Dot-matrix comparisons may also be used to identify repeats within a single sequence. If a sequence is compared to itself, in addition to the uninterrupted diagonal line of self-identity across the matrix, any direct repeats within it are seen as additional short diagonals, displaced to either side of the main diagonal. Similarly, if a nucleic acid sequence is compared with its inverse complement, any inverted repeats appear as short diagonals, generally lying off the position where a central diagonal is. (This can be seen if one considers a dot matrix comparison between a sequence 5′-TTAAGCxxxGCTTAAxxx-3′ and its inverse complement, 5′-xxxTTAAGCxxxGCTTAA-3′.)

13.2.4 Multiple Sequence Alignment

We have discussed the problem of alignment of two sequences and its solution by dynamic programming. What happens when we want to find the best alignment between a group of related sequences? Unfortunately, although dynamic programming can, in theory, find the best alignment, the time complexity is n^x, where x is the number of sequences to be compared. For $x = 3$ this is just about feasible if one is in no hurry; for more than three sequences, dynamic programming is not viable.

Because of this, heuristics are employed in the programs currently in use, one of which, Clustal [6], we consider here (Figure 13.2(c)). At the time of writing, ClustalX is the latest version that can be downloaded from www.clustal.org/, although a more advanced version can be used online at the EBI website (www.ebi.ac.uk/Tools/msa/clustalo/). Another popular program, not discussed here, is T-coffee, available at www.tcoffee.org/ [7].

Clustal's heuristics have evolved over time, but it is useful to consider one aspect of them here. The assumption made is that regions of similarity between the most closely related sequences are also represented in the multiple alignment (albeit interrupted by gaps), and that it is valid to use the alignment of these sequences as a nucleus from which to build the multiple alignment. Initially a series of all possible pairwise alignments is performed, generating a guide tree of the similarities between the sequences. The highest scoring alignment is presumed to represent the similarity between the two most related sequences and is used as the starting point.

The preliminary alignments use a standard amino acid–scoring matrix (see Figure 13.1), but subsequent comparisons use dynamically generated position-specific matrices based on the alignment's current state at any time. Consider, for example, the initial alignment of two sequences in which at position x a D (aspartic acid) is aligned with an E (glutamic acid), whereas at position *y* a D is aligned with another D. The revised scoring matrix used to align the sequence most closely related to the first two (as determined by the guide tree) assigns a lower score for matching a D to position x than for matching it to position y, and indeed the score for a match at position y is higher than in the original standard scoring matrix. Conversely, the score for aligning an E at position *y* is lower than in the original scoring matrix.

There are a number of problems with this approach, which have been the stimulus for continual improvements in the program. One problem is that the alignment can easily be biased if there are several very similar sequences. Another is that an initial gap insertion, which can clearly be seen to be erroneous after other sequences have been added, becomes set in stone. A similar sort of problem occurs where related proteins have core regions of similarity but very different N- or C-terminal sequences. After an initial run, users may therefore have to remove certain sequences or remove, say, the N-terminal regions of all sequences in order to obtain a satisfactory multiple-sequence alignment.

13.2.5 Database Searching

After obtaining a new sequence, one frequently wishes to know whether it is related to any previously determined sequences. This involves successive pairwise comparisons with sequences in a database. Although the time complexity of an individual comparison may only be n^2, the number of sequences in the publicly accessible GenBank (www.ncbi.nlm.nih.gov/genbank/) currently exceeds 100 million. Thus, a heuristic approach is needed to search this completely.

The premier program (or suite of programs) for this purpose is BLAST [8], which is most frequently used on the NCBI website (http://blast.ncbi.nlm.nih.gov/), although it can be downloaded and installed locally by those with the appropriate technical expertise. BLAST is continually evolving and employs a variety of devices. For this reason, we mention only one aspect of the heuristic approaches it uses. The assumption made is that any good alignment between the query and the target sequences contains short high-scoring ungapped matching regions (segments).

The program starts by indexing the segments in the query sequence, which allows rapid identification of any small high-scoring diagonal segments in the comparison (a BLOSSUM scoring matrix is used for proteins). It then determines whether these can be extended diagonally to encompass other such segments. If the search is unsuccessful, it is abandoned and the next sequence is considered. Time is allocated only to more extensive comparison of the highest-scoring target sequences of the initial sweep.

The output from BLAST shows a pairwise dynamic programming alignment with the highest-scoring targets, and this generally includes gaps. However, it should be stressed that the heuristic described allows BLAST to look for related sequences only on the basis of ungapped regions of similarity. This was the situation in initial versions of BLAST, although a "gapped BLAST" has been introduced.

A variant of BLAST, PSI-BLAST [9], performs "profile searching." Rather than searching with a single-query sequence, GenBank is searched with a profile derived from a family of similar sequences in order to discover additional, perhaps more distantly related, members of the same family. This is similar to multiple-sequence alignment, where the final scoring matrix used to make a multiple alignment can be considered as the scoring matrix for a profile of aligned sequences.

The beauty of the BLAST web interface is that it is extremely clean and simple to use. A query sequence can be dropped into a form field, a button pressed, and a result generally obtained within a minute. However, many additional options are available if this simple approach produces unsatisfactory results, and the website contains clear and extensive information about BLAST, which must be the most widely used bioinformatics program on the Internet.

13.3 MOLECULAR EVOLUTION AND PHYLOGENETICS

Molecular evolution is the study of the evolution of certain biological macromolecules: DNA, RNA, and proteins. The two primary phenomena assessed are the rate of change over generations and the pattern of the differences between or within specific organisms. The changes that occur can be point mutations (substitution and deletion or insertion of a single base) or large events such as inversions, translocations, and larger insertions/deletions. Studying molecular evolution allows us to discern mechanisms for evolutionary changes and to reconstruct evolutionary history.

Bearing in mind the "central dogma" of molecular biology, that *DNA makes RNA makes protein* [10], it might seem redundant to mention all three of these macromolecules in our definition of molecular evolution, however, it should be noted that many viruses have RNA as their genetic material and, in fact, their replication has a particularly high rate of mutation: between 10^{-3} and 10^{-5} per base per generation [11]. If genetic changes do not result in a codon that codes for a different amino acid, the mutation is said to be *silent*. For some purposes, then, it may be more prudent to compare amino acid sequences rather than genetic sequences.

To make the distinction between change within and between organisms, it is important to introduce the concepts of homology, orthology, and paralogy. Homologous sequences share an ancestry and have deviated from each other by either a speciation event (orthology) or a gene duplication event (paralogy). Homology is *inferred from observations of similarity*, but is distinct from similarity

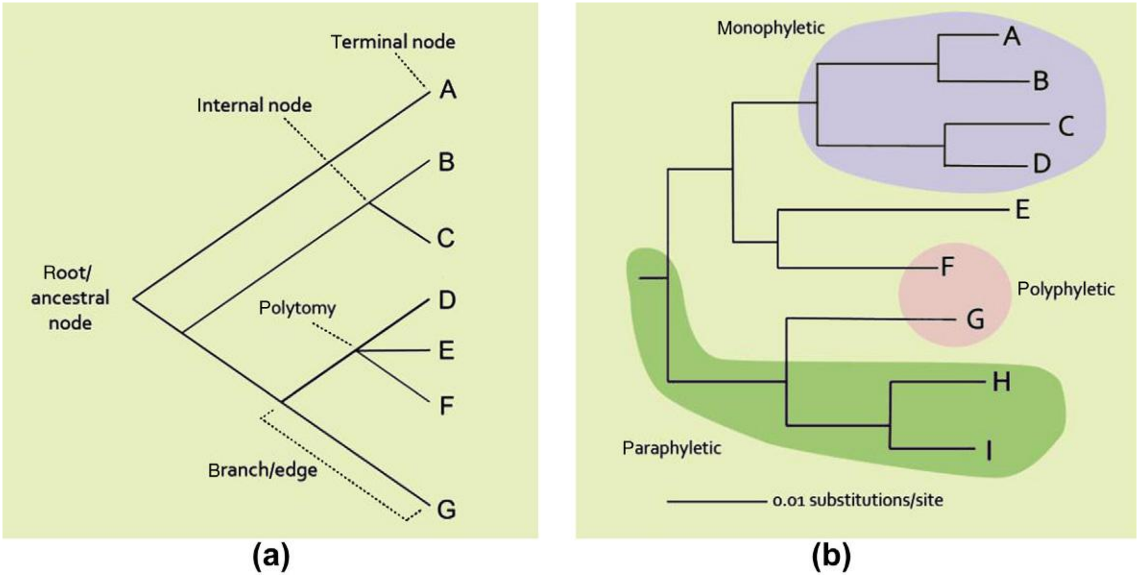


FIGURE 13.3 **Phylogenetic relationships using various types of tree.** (a) Common features of a cladogram. The terminal nodes A–G are the sequences used to infer the organization of the tree itself and to infer internal nodes. The root of the tree is the node that is ancestral to all others. The polytomy node has a degree of >3 (i.e., it relates directly to four or more other nodes). Note that branches are also called edges and nodes are also called vertices. (b) Monophyletic, polyphyletic, and paraphyletic groups. Monophyletic groups (clades) must by definition include all child nodes of any of its members. Paraphyletic groups, in contrast, must contain all ancestors of all nodes in the group but not all child nodes. Polyphyletic groups are composed of unrelated nodes descended from more than one ancestor. *Source: Hand-drawn by Emma Fitzgibbon based on example provided by author.*

in that the sequences must be related *by descent*. Often the term *homology* is used either as a synonym for similarity or with the assumption that it is necessarily implied by similarity, although strictly speaking this is not the case. Sequences can be similar but not homologous as a result of convergent evolution, and they can be dissimilar yet homologous, although this is more difficult to ascertain.

Phylogenetics deals with identifying evolutionary relationships between organisms or between genes. In the case of the evolutionary relationships between organisms, this is also often known as "phylogenetic systematics." Using molecular rather than morphological data has revolutionized systematics, from redefining our place among the primates [12] to allowing us to analyze the evolution of viruses and even revealing archaea, an entirely new domain of life [13]. Phylogenetics is also extremely useful in understanding the spread of contagious disease [14] and the evolution of the responsible pathogens [15].

13.3.1 Tree Terminology

Multiple sequence alignments, as explained in Section 13.2.4, help identify homology and reconstruct evolutionary history. Alternatively, it can be said that variation between sequences is used to infer phylogeny. We depict phylogenetic relationships using various types of trees (Figure 13.3). Evolutionary trees are graphs consisting of nodes and branches, which in mathematics are often referred to as vertices and edges. Trees may be either rooted or unrooted, where the root is a node that depicts an ancestor common to all of the other nodes in the graph. Intuitively, then, rooted trees depict directionality in evolution, whereas unrooted trees merely show evolutionary distances. Rooted trees are commonly drawn with the root on the left and the terminal nodes (or tips)—for which we have sequence data—on the right. In between the root and the terminal nodes, there may be internal nodes that depict hypothetical genes or organisms—for which we have no experimental data.

The lengths of the lines that constitute the branches can mean different things depending on the type of tree. Cladograms (Figure 13.3(a)) show only the orders of branching; the lengths of the branches are meaningless. In phylograms (as shown in Figure 13.3(b) and also known as additive trees), the horizontal branch lengths correspond to evolutionary changes whereas the vertical distances are meaningless and only used for separation and clarity. Ultrametric trees (also known as dendrograms) are similar to phylograms, but the root must be equidistant to each terminal node. In this case, the distance represents either actual time or "time" on a molecular clock. The concept of a molecular clock is that sequences evolve at an approximately constant rate, and so the genetic difference between two species or genes is proportional to the amount of time since the speciation or gene duplication event. It should be noted, however, that there is no one universal molecular clock—differences between genomes of different species, and throughout different genomic regions of the same species, lead to different rates of mutation/evolution.

It is often useful to describe groups of nodes and how they relate to one another. Clades are groups of nodes united by ancestry, in which every descendant of every node is present; in other words, a clade is an entire branch. This must not be confused with grades, which are nodes grouped by characteristics such as morphology. Clades are also described as monophyletic groups. Two concepts that contrast with monophyly are paraphyly and polyphyly (Figure 13.3(b)). Paraphyletic groups are made up of consecutive nodes that share common ancestry but in which not all descendents of all nodes are present. Polyphyletic groups do not include a common ancestor of all of its members. An alternative to cladistic classification is phenetic classification, which is based on overall sequence similarity rather than ancestry. Phenetic classification particularly causes different classifications when different rates of evolution occur in different branches of the tree.

Tree diagrams represent genes, species, or both. They therefore can describe paralogy, orthology, or the two of them. In other words, depending on the type of information in the diagram, a branching can represent either a gene duplication event or a speciation event. Trees are similar to hanging mobiles (as in kinetic art) in that the branches can be rotated without altering the meaning of the tree or the relationships between terminal nodes. The specific order in which the terminal nodes are represented is merely stylistic and not inherently meaningful.

Trees allow us to distinguish between ancestral and "derived" (newly and separately occurring) character states or sequence variations. A plesiomorphy is an ancestral character state, whereas an apomorphy is a derived state. If an apomorphy occurs in one sequence only, it is described as autapomorphic, whereas if it occurs in several sequences it is described as synapomorphic. Homology describes two identical character states in multiple terminal nodes inherited from a common ancestor. In contrast, homoplasy describes two identical character states in multiple terminal nodes that have occurred independently.

The degree of a node is the number of branches that are connected to it. A tree is said to be fully resolved if the root has only two children and none of the internal nodes has a degree of more than three—that is, one parent node (ancestor) and two child nodes (descendants). A node having a greater degree than this is described as a polytomy (Figure 13.3(a)). If a tree has no polytomies, it is said to be fully resolved, whereas a tree completely unresolved is said to be a "star tree." Polytomies are described as "hard" if they indicate events of simultaneous divergence for all descendants involved. They may be described as "soft" if they represent an uncertainty in the order of divergence. Other related terms here are bifurcation (divergence into two branches) and trifurcation (divergence into three sequences).

13.3.2 Principles of Tree Inference

Alignments of terminal node sequences are used to make inferences about the relationship of the nodes and about the inferred ancestral nodes for which there are no data. Different parts of the genome are subject to different mutation rates because of factors such as GC content or the presence of CpG islands [16]. Genetic changes are conserved to different extents depending on the type of region they are found in. For example, one expects a sequence coding a binding domain of a protein to be highly conserved. Whether a mutation is silent or not further indicates the possibility of it "surviving" the process of evolution. These factors are important because, if sequences become saturated with polymorphisms, they become essentially random and no longer useful for this type of analysis. There are two main classifications: (1) it can be either "distance" or "discrete," or (2) it can be "exact" or "heuristic" (Table 13.1). Each combination can have a multitude of specific algorithms and implementations.

Discrete methods consider the differences in nucleotides directly, whereas distance methods first convert the data into a score for each pair of sequences based on the number of differences present. These dissimilarities can be partly shown by a parsimony tree (Figure 13.4(a)) and a distance tree (Figure 13.4(b)), which are identical except that the former shows additional information. A criticism of distance methods is that

TABLE 13.1 Four Main Branches of Tree Inference, as Combinations of Distance/Discrete and Exact/Heuristic Approaches

	Distance	Discrete
Exact	Minimum evolution	Maximum parsimony Maximum likelihood
Heuristic	UPGMA Neighbor joining	Bayesian inference

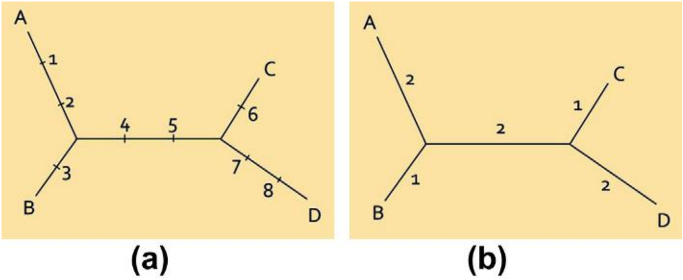


FIGURE 13.4 **Parsimony tree and distance tree.** (a) Parsimony tree. A–D represent sequences and numbers 1–8 represent variations in sequences. (b) Distance tree, identical to the parsimony tree but showing "distances" for each branch (i.e., the number of variations). *Acknowledgment: Hand-drawn by Emma Fitzgibbon based on example provided by the author.*

they are phenetic rather than cladistic. Another more subtle difference between the two, in how the data are treated, is whether to include or ignore insertions and deletions.

The other main classification of a tree-building method is whether it is exact or heuristic. Exact methods can be referred to as scoring or optimality, and some heuristics can be referred to as clustering. The aim of an exact method is to assemble all possible trees and score them based on a chosen optimality criterion. Clustering starts with a tree of three terminal nodes (only one unrooted tree of this size is possible); with each iteration of a process one additional node and branch are added until all sequences have been mapped to a final tree.

The advantage of clustering is its high interpretability, having only one tree as a result, but this is also its weakness as several trees may be almost as likely, based on the data, but will go completely unreported. In contrast, optimality/scoring presents all options with their associated scores to be ranked and appropriately interpreted. Clustering is appropriate for large amounts of data, particularly a large number of sequences rather than length of sequences. In these cases, generating and scoring all hypothetical trees is simply too computationally demanding. With five sequences, there are 15 possible trees; however, a meager increase to 10 sequences results in 2,027,025 possible trees (Table 13.2). These formulas, which describe how many trees are possible, are as follows:

1. Number of unrooted trees $=(2n-5)!!$, where "!!" denotes a double factorial. The double factorial of an odd number is the product of all positive odd integers up to and including the input itself; for example. $(2n-5)!!=(2*5-5)!!=5!!=1*3*5=3*5=15$, where $n=5$. Similarly, the double factorial of an even number is the product of all even integers up to and including the input. Including 1 in the case of an odd double factorial is, of course, mathematically equivalent to excluding it.
2. Number of rooted trees $=(2n-3)!!$

13.3.3 Methods of Tree Inference

There are a multitude of methods for tree inference, and it may be difficult to decide which is the best for, or the most appropriate, to the data, especially when there are an intermediate number of sequences. David Penny and colleagues suggest five criteria for assessing an algorithm or program [17]. Criterion 1 is that it is efficient (i.e., fast). As shown in the previous section, the increase in the number of possible

TABLE 13.2 Possible Tree Configurations

Sequences (n)	2n – 5	Unrooted Trees (2n – 5)!!
3	1	1
4	3	3
5	5	15
6	7	105
7	9	945
8	11	10,395
9	13	13,5135
10	15	2,027,025

Note: Given an increasing number of sequences, the increase in the number of configurations is given by $(2n-5)!!$.

trees with a given number of sequences is exponential, so speed is of great importance. If the number of sequences is too great, a global optimum cannot be found, a problem described as NP-complete [18]. NP-complete problems are grouped by the criterion that if one problem can be efficiently solved they all can, and that no efficient algorithm is known (and likely will never be known).

The second criterion is that the method is powerful—it does not require large numbers of data to come to a conclusion. With shorter sequence lengths, sampling errors are more probable, and as the length of a sequence is increased, the method may select different trees, converging on a single final tree at a particular sequence length. The smaller the length required for this convergence, the more powerful the method is.

The remaining criteria are consistency, robustness, and falsifiability. A model is consistent if, given ample data, it converges on the "correct" tree based on one's interpretation or model of molecular evolution. Robustness concerns whether the method is still accurate even when its assumptions are slightly violated. This can be tested by simulations on calculated datasets. Falsifiability is a general criterion for any scientific method, although many tree inference methods do not meet it. On a related note, one should be concerned about any method that does not indicate when its assumptions are violated (a clustering method will select one tree whether or not assumptions are violated). It may also be worth assessing scalability, that is, whether the method is accurate with both large and small datasets.

Maximum parsimony describes the adoption of the simplest interpretation of a set of data, which, in general, is often referred to as "Occam's razor" after the philosopher William of Ockham. Maximum parsimony is an exact method that uses discrete data. It selects the tree, of all possible trees, that requires the fewest evolutionary changes across all sequences being studied. It is not estimated where along a branch a change might have happened, only which change and on which branch. Maximum parsimony's benefits are that it is simple and easy to interpret and all trees may be scored and represented. It relies on the assumption that evolutionary change is rare and seeks to explain the variations present in the data by proposing the smallest number of mutational events.

Transitions are the replacement of a pyrimidine with a pyrimidine or a purine for a purine, whereas transversions are the replacement of a pyrimidine with a purine or vice versa. Transitions are substantially more common [19]. The simplest implementation of maximum parsimony is a model where transitions and transversions are equally likely and where the rate of evolution is consistent throughout the sequence length.

A more accurate model would assign a higher score (greater penalty) to transversions than to transitions. As mentioned earlier, different genomic regions are more prone to mutation and conservation of ancestral sequence, giving a varied rate of change throughout the genome. Another way that the method may be slightly altered to improve its approximation to actual events is to weight different sequence regions. In other words, different regions are scaled differently to represent the likelihood of a mutation in that particular region.

"Long branch attraction," first described by Joseph Felsenstein [20], is an example of a situation where maximum parsimony is prone to fail in selecting the correct tree, and the length of the branch represents the number of substitutions that occur along the branch. This is characterized by a small number of sequences that are either quite divergent or show highly varied rates of evolution; hence, some branches are much longer than others. In this situation, long branches tend to be grouped together because of an increased likelihood that the same substitutions are occurring along the branches. These corresponding homoplasies are being interpreted as homologies under maximum parsimony. They are much more likely to occur in smaller trees, where the lineages are relatively similar to begin with, than in distant branches of larger trees.

Where probability describes how often an event will occur out of a known set of all possible outcomes, likelihoods are relative to true probabilities and compare how likely are two or more events out of a set of events of unknown size. For example, a probability for rolling a 6 with a fair six-sided die is 1/6. Supposing the die has an unknown number of sides but we know 6 occurs twice as often as 5; then we cannot assign probabilities, but likelihood can still describe the *relative* probability of rolling a 5 compared to rolling a 6. Maximum likelihood [21] aims to associate a likelihood score to each tree—the likelihood that the input data have been produced by the tree. More formally it can be described by $\mathbf{L_D = Pr(D|}\mathrm{H})$, where the likelihood of the data occurring, $\mathrm{L_{D,}}$ is calculated by the probability of the data, D, given the hypothesis, H (the phylogenetic tree being assessed).

Maximum likelihood involves an explicit model that represents the probabilities of a position in the sequence changing from any base to any base (or the base remaining the same). Also required for maximum likelihood are, of course, the sequences and the proposed tree. Often the natural log of the likelihood, rather than the direct likelihood, is calculated. This is more prudent, seeing as the likelihoods are often small (the more data, the smaller they are). Using the model described, probabilities for the content of each nucleotide position can be generated (e.g., the likelihood that position 5 of the sequence is A, A, A, and G across four different sequences), given the model and the distances between the sequences on the hypothesized tree.

To generate the log-likelihood for a branch, the position-specific log-likelihoods are summed. The branch lengths are then optimized by projecting likelihoods for a variety of branch lengths and selecting the maximum. The maximum likelihood for the tree is then generated from these branch-specific likelihoods. One of the advantages of maximum likelihood is that the likelihood of the data occurring given a specific tree can be tested against the null hypothesis. Maximum likelihood allows a more detailed model than maximum parsimony, but it suffers even more with respect to computational time and large datasets.

Minimum evolution is similar to maximum parsimony but differs in that, rather than fitting and counting the individual nucleotide changes to a tree, it uses pairwise distances between sequences (a count of the number of nucleotide differences). The total length of the tree—that is, the sum of all branch lengths—is computed for each possible topology and the one with the smallest total length is selected. A disadvantage of this method is that, while it informs topology and evolutionary distance between sequences, it does not inform on specific base changes. Least squares methods are used to calculate the branch lengths based on the differences between sequences.

Clustering methods (which are heuristics) are much faster than exhaustive optimality criterion approaches, and so they are more appropriate for large datasets. Neighbor joining starts with a star tree. A pairwise distance matrix is calculated for the sequences and indicates the closest sequences to join with a new internal node. New internal nodes are iteratively added in this manner until the tree is fully resolved. Unweighted pair group method with arithmetic means (UPGMA) creates an ultrametric tree and relies on the assumption of a constant rate of evolution. Distance clustering methods such as these can be used to find a close approximation to the desired tree; exact methods can be used to optimize the tree based on a particular optimality criterion such as minimum evolution [22].

In summary, maximum parsimony, maximum likelihood, and minimum evolution are good choices when a small number of sequences are involved, or when one wishes to investigate a given set of phylogenies rather than infer them. Maximum likelihood allows us to more fully model evolution, so it is more powerful, although it requires knowledge of an appropriate model and more assumptions must be met. Minimum evolution is similar to maximum parsimony except that it assumes that the tree with the least evolutionary distance is the correct one. While these methods all perform very well with a trivial number of sequences, when a larger number of sequences is involved they become too computationally expensive. One must therefore use a heuristic approach.

By employing these methods, we are in effect estimating the true population using the sample that was sequenced. Bootstrapping involves resampling the sample population and reapplying the tree-finding method on each pseudoreplicate of the data. By doing this, counts for each tree can be generated and the ratio of a count for a tree to the total number of pseudoreplicate-generated trees can be generated as a score. These scores can be used as a measure of the method's precision, giving confidence that the selected tree is robust to sampling error, although it does not necessarily speak to the correctness of the prediction (e.g., because of an incorrect model or an unusual phylogeny). Felsenstein suggested that bootstrap values of >95 can be interpreted as being well supported. Bootstrap values are also often applied to each node of a tree.

A relatively new approach to tree building involves using Bayesian inference. Algorithms employing Bayesian inference are heuristics and can be applied with detailed models of evolution, as in maximum likelihood [23]. In contrast to the other methods outlined, bootstrapping is not required because the output expresses the uncertainty of each tree based on a probability distribution.

There are many programs that can be used for multiple-sequence alignment, tree inference, and preparing a diagram for presentation. A free tool that does all of these and uses a variety of tree inference methods (including all of those mentioned previously) is molecular genetics evolutionary analysis (MEGA). This tool is noncommercial, has a graphical user interface (GUI), and is available for Mac and Windows [24]. MrBayes is also worth mentioning as it can take advantage of modern graphics cards and their GPUs, which often now contain hundreds of cores [25], allowing for a highly parallelized version of a Bayesian inference algorithm that increases the speed of inference.

13.4 MOLECULAR GRAPHICS AND PROTEIN MODELING

Biological molecules such as proteins are complex structures, and representing their interpretable illustrations requires specialized techniques. Molecular graphics provides techniques to visualize and manipulate molecules on a computer. It can be used to give a general impression of a protein folding pattern; map residues that are assumed to be involved in a protein's function to a 3D representation of the protein to isolate an active site; classify and compare protein folding patterns; analyze changes between individual related structures or different conformational states of a single molecule; study the interaction of a ligand (small molecule) with a particular protein for drug development; and design and model a new structure. Because protein structure forms the main focus of molecular graphics, we discuss the basics of it in the following subsection.

13.4.1 Protein Stability

Proteins must fold to perform their intended biological function. They fold spontaneously to their active native 3D structure, which is determined by their corresponding sequence of amino acids. To form this native structure, a protein should optimize the interactions between its residues because each individual conformation of the main chain puts different residues in contact with each other.

The stabilities of different conformations are defined by the interactions of the side chains and main chain with each other and with the solvent environment of the cell. Moreover, the steric restriction on side-chain mobility is important for conformation stability.

13.4.1.1 Allowed Main-Chain Conformation

The main-chain conformation of each non-glycine amino acid is limited to two conformational states. This limitation comes from the angle of rotation around the peptide bond between two residues and the torsion angles between N–C_α and C_α–C single bonds of all residues. The torsion angle around the peptide bond is called ω while the torsion angles between N–C_α and C_α–C single bonds are called ϕ and ψ angles, respectively. Figure 13.5 shows the ϕ and ψ angles on a fragment of a polypeptide. The sequence of the ϕ, ψ, and ω angles define the conformation of the protein's main chain.

13.4.1.2 Permitted Torsion Angles

The peptide bond is planar and has two states: *trans*, $\omega \approx 180^\circ$, and *cis*, $\omega \approx 0^\circ$. In the *trans* configuration, the two alpha carbon atoms of the connected amino acids are on the opposite sides of the peptide bond, whereas in *cis* configuration they are on the same side of the peptide bond. In most cases, the peptide bonds in proteins are *trans*. This preference can be explained by the steric clashes that occur between groups attached to the alpha carbon atoms in *cis* form, which hinder formation of this configuration.

G.N. Ramachandran recognized that steric collisions between atoms prevent some combination of ϕ and ψ angles and, for the *trans* configuration, ranges of *ϕ* and *ψ* angles fall into defined regions in a graph called the Ramachandran plot (Figure 13.6) [26]. According to permitted ϕ, ψ, and ω angles, preferred conformations of the main chain lead to recurrent structures in proteins—namely, alpha helix, beta sheets, and turns.

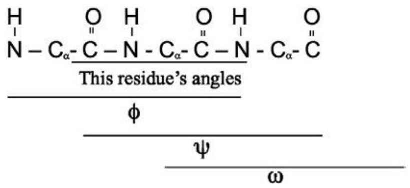


FIGURE 13.5 Conformation of each amino acid in a protein is restricted by rotation around two single bonds. Phi (Φ) is the torsion angle around the single bond between the nitrogen and the alpha carbon atom in an amino acid, while Psi (Ψ) is the rotation angle around the single bond between the alpha carbon atom and the carbonyl carbon atom. *Source: James Milner-White (2011), "Molecular Graphics," Lecture/Class, University of Glasgow, unpublished.*

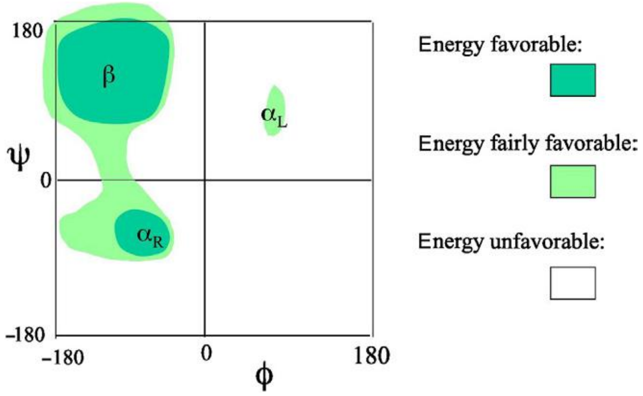


FIGURE 13.6 Ramachandran plot illustrating the allowed combination of Φ and Ψ torsion angles. The principle that two atoms cannot be in same place at the same time leads to exclusion of almost three-quarters of the plot. *Source: Adapted from Milner-White (2011). "Molecular Graphics," lecture, University of Glasgow.*

13.4.1.3 Alpha Helices

The alpha helix is a rod-like structure whose inner section is formed by a tightly coiled main chain, with its side chains extending outward in a helical array. The alpha helix structure takes advantage of the hydrogen bond between CO and NH groups of the main chain to stabilize. The CO group of each amino acid forms a hydrogen bond with the NH group of amino acid four residues earlier in the sequence. Therefore, in an alpha helix, all main-chain CO and NH groups are hydrogen bonded except in those amino acids close to the end of the helix.

The "screw sense" of an alpha helix can be right-handed (clockwise) or left-handed (counter-clockwise). Despite the fact that, based on the Ramachandran plot, both right-handed and left-handed alpha helices are among the permitted conformations, the right-handed alpha helix is energetically more favorable because of fewer steric clashes between the side chains and the main chain. Thus, all alpha helices in proteins are right-handed.

13.4.1.4 Beta Sheets

The beta pleated sheet—in short, the beta sheet—is another recurrent structure in proteins; it is noticeably different from the alpha helix. In a beta sheet, the main chain, called the beta strand, is fully extended rather than tightly coiled, and the side chains of adjacent amino acids point in opposite directions. A beta sheet consists of two or more beta strands linked by hydrogen bonds. Adjacent strands that form the beta sheet can either run in opposite directions (antiparallel beta sheet) or in the same direction (parallel beta sheet).

In antiparallel conformation, the CO and NH groups of each amino acid are hydrogen-bonded to the NH and

CO groups of the amino acid located in alignment on the adjacent strand. In parallel conformation, the NH group of each amino acid is hydrogen-bonded to the CO group of the amino acid on one side of the adjacent strand's inline amino, while the CO group is hydrogen-bonded to the NH group of the amino acid located on the other side of the adjacent strand's inline amino acid. Typically, beta sheets consist of 4 or 5 beta strands where they can be all antiparallel, parallel, or mixed. Beta sheets can be relatively flat or to some extent twisted, which makes them more structurally diverse than alpha helices.

13.4.1.5 Turns and Loops

Proteins typically have compact, globular shapes formed by combination of alpha helices and beta sheets. However, to obtain these compact shapes, they require reversals in the direction of their chain. The reverse turn, also known as the beta turn or hairpin bend, provides a common structure to satisfy the chain reversal requirement. To stabilize this structure against sudden changes in direction, the reverse turn forms a hydrogen bond between the CO group of the "i" residue and the NH group of the "i+3" residue of a polypeptide chain.

Loops are another type of structure responsible for chain reversals and are more intricate than reverse turns. Although they do not have periodic structures similar to alpha helices and beta sheets, they are rigid and well defined most of the time. Reverse turns and loops always lie on the protein's surface and hence often participate in the interaction of the protein and other molecules.

The three-dimensional structures of proteins, which consist of different combinations of the recurrent structures, are stored in the Protein Data Bank (PDB) repository as text files in the .pdb file format.

13.4.2 Protein Data Bank File Format

The pdb file format provides the atomic coordinates for the main-chain and side-chain atoms, secondary structure assignments, and atomic connectivity. Pdb files may provide only the asymmetric unit—the smallest portion of a molecule that, upon symmetric operations such as rotation and translation, can be used to assemble the whole molecule. Asymmetric units are usually used by crystallographers to refine the coordinates of the structure against the experimental data.

A typical pdb file has a 4-digit code of the form "1xyz" and contains the following sections: title, primary structure, heterogen, secondary structure, connectivity annotation, miscellaneous features, coordinate, and connectivity [27].

Title. This section describes the experiment and the biological macromolecules present in the file.

Primary structure. This section contains the sequence of residues in each chain of the protein. Furthermore, it accommodates the chain identifiers and sequence numbers used by other records to link to this section.

Heterogen. Any nonstandard residues such as ions, prosthetic groups, and solvent molecules are described in this section. Each nonstandard residue is assigned the sequence number, chain identifier, and number of that type of residue described in the coordinate section.

Secondary structure. This section describes the helices and sheets that are distributed along a protein chain. Helices are named, numbered, and classified by type. Moreover, the total length and the start and end residues of the helix structure are noted. Sheets are described using the same factors.

Connectivity annotation. This section specifies the existence and location of linkages between residues which are not directly expressed by the protein's sequence of amino acid of the protein.

Miscellaneous features. Properties such as environments surrounding a nonstandard residue or the assembly of the active site may be provided in this section.

Coordinate. This section consists of atomic coordinates for both standard residues and nonstandard chemicals such as water molecules or atoms presented in Heterogen section.

Connectivity. Information regarding the atomic connectivity between atoms, for which coordinates are provided in coordinate section, can be retrieved from this section.

Molecular visualization systems such as PyMOL [28] or Jmol [29] can be employed to view the three-dimensional structure of proteins described in pdb files. The Java applet version of Jmol is particularly useful as it may be embedded into web pages; however, PyMOL is more extensive in terms of features. Both are open-source visualization tools that show structures using various standard representation styles (Figure 13.7) and output them as either images or video files.

Two standard representations are ball-and-stick and space-fill. Ball-and-stick represents atoms by appropriately colored balls and the bonds between them as lines or "sticks." This is useful for visualizing the geometry of the bonds between atoms, but it gives no indication of the space occupied by the atoms. Space-fill, on the other hand, shows the space occupied by the atoms (since each atom is represented by a sphere with the radius equal to its van der Waals radius), which allows the user to more easily identify steric clashes.

Another useful representation is one in which beta sheets and alpha helices are clearly shown by cartoons and ribbons. The cartoon model simplifies protein structure to illustrate the direction of the protein chain, protein folding, the element of the secondary structures, turns and loops,

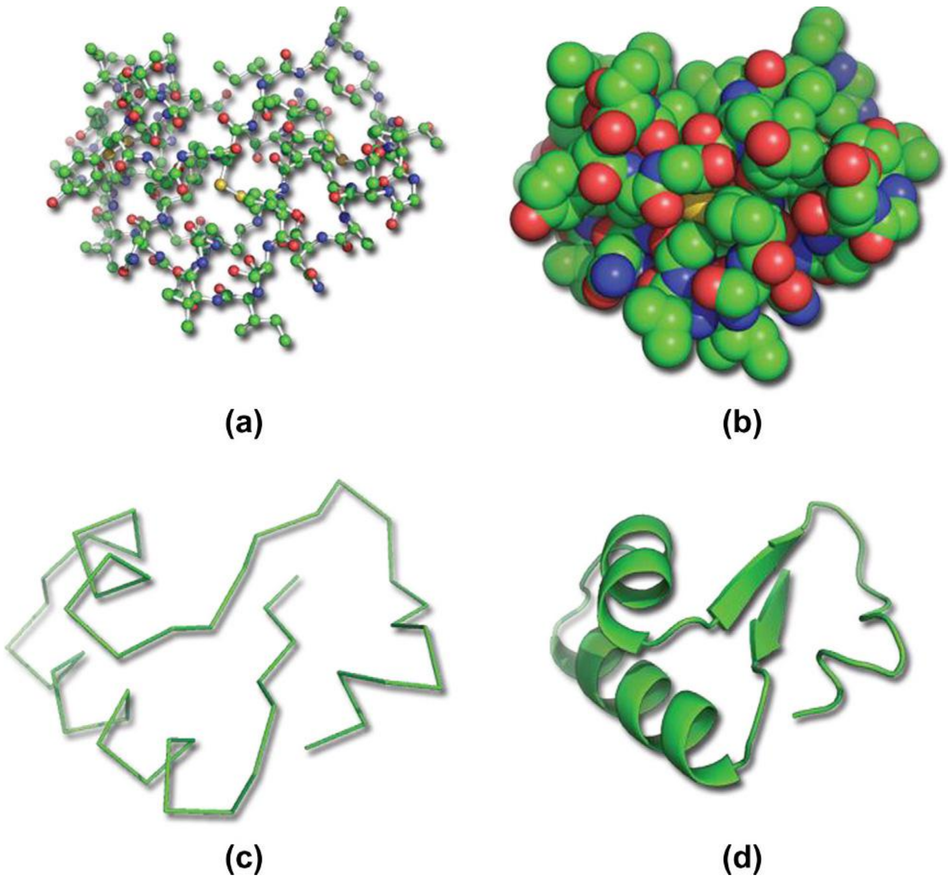


FIGURE 13.7 **Various representation methods as depicted by PyMOL.** The images show (a) ball-and-stick, (b) spacefill, (c) ribbon, and (d) cartoon.

and domain structures. The ribbon model can be categorized as a cartoon model, but the algorithm used to generate ribbon representation is different. The ribbon model employs the B-spline curve to render proteins, whereas the cartoon model is derived from models of protein secondary structures hand-drawn by Richardson [30].

13.4.3 Protein Modeling

Despite the increasing success of approaches for obtaining protein sequences, for the vast majority of known sequences, no structural information is available. The goal in protein modeling is to predict the protein's structure from its sequence with a level of accuracy comparable to the result acquired experimentally. This can fill the gap in structural knowledge in fields such as structure-based drug design, which otherwise must rely only on experimentally derived structures. Moreover, protein modeling is the only way to acquire some knowledge of protein structure, when experimental methods fail. For instance, many proteins are too large for NMR analysis or cannot be easily crystalized for X-ray diffraction methods. Methods for predicting the 3D structure of proteins include homology modeling, fold recognition, and de novo structure prediction.

13.4.4 Homology Modeling

Among approaches to predicting 3D structures, homology modeling can be considered as the easiest because of two observations. First, the structure is determined by the 1D sequence of the protein's amino acids [31]. Thus, the amino acid sequence should contain enough information to obtain the structure. Second, as proteins evolve their structure and sequence change. However, the rate of modification is much slower in structures than in sequences [32,33]. This observation is summarized in Figure 13.8.

Based on the figure, if the percentage sequence identity of two sequences is above the threshold, it is likely that they fold into approximately the same structure. If, on the other hand, the percentage sequence identity falls below the threshold, it is not evident whether two sequences fold into similar or different structures.

The sequence with a known structure is called the template, while the sequence with an unknown structure is called the model. Simple similarity search tools such as

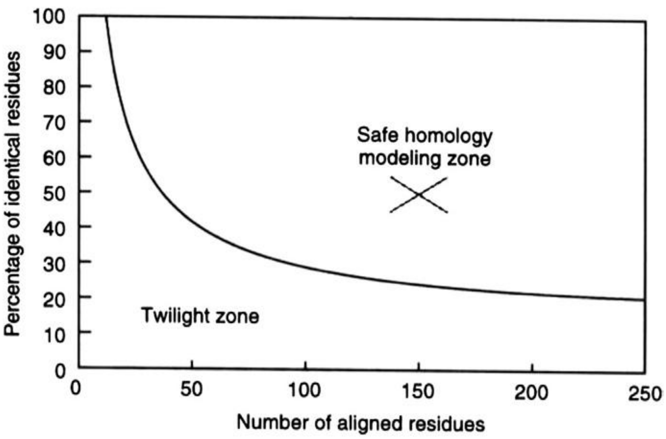


FIGURE 13.8 **Relationship between percentage sequence identity and number of aligned residues.** The zone above the threshold is called the "safe zone"; the zone below is called the "twilight zone." For longer sequences, the percentage identity that falls in the safe zone can be less than 25%. *Source: Adapted from Sander et al.* [31].

BLAST (local alignment) [7] and FASTA (global alignment) [34] can determine the identity between the template and the model. Once the percentage identity in the aligned region of the template sequence and the model sequence falls into the safe zone shown in Figure 13.8, a model can be built.

Homology modeling (Figure 13.9) is a multistep procedure, including the following steps: template recognition and initial template–model sequence alignment; alignment correction; backbone generation; loop modeling; side-chain modeling; model optimization; and, finally, model validation.

13.4.4.1 Template Recognition and Initial Alignment

In the safe homology modeling zone, the percentage identity of the model and the template sequence is high enough to be detected with BLAST or FASTA. To identify the template sequence, either program compares the query sequence with all sequences of known structure in the Protein Data Bank repository. It is possible to use the sequence with the highest sequence identity as the template sequence or to combine multiple sequences into one template sequence. The latter method is used by the Swiss-Model [35] and Robetta servers [36].

13.4.4.2 Alignment Correction

After retrieving the template sequence, more sophisticated methods can improve the performance of the final model. For instance, in regions where the percentage sequence identity is very low, it might be difficult to align the template and model sequences. One solution is to increase the identity percentage via other sequences from other homologous proteins, which are found by a multiple-sequence alignment, and use implicit structural information from the alignment. For example, if at certain positions of the sequences only hydrophobic residues are found, it is likely that these residues are located at the hydrophobic core and are therefore "buried." Position-specific scoring matrices, also called profiles [37,38], consider these features during the alignment procedure; programs, such as MUSCLE [39] and T-Coffee [7], use these profiles to improve multiple-sequence alignment.

13.4.4.3 Backbone Generation

Once the alignment is ready, the model-building process starts off with backbone generation. In the case of a single template, this is trivial, and one can simply transfer the backbone coordinates of aligned residues and the side-chain coordinates of conserved residues. In the case of multiple templates, an average structure is used where individual template contributions are weighted by the local degree of sequence identity.

13.4.4.4 Loop Modeling

As described in Section 13.2.4, sequence alignments may contain gaps; the same also occurs with model and template alignment. These gaps can be easily addressed by omitting the residues from the template sequence. Similarly, inserting residues into the model sequence addresses the gaps in the template sequence. Both cases therefore lead to conformational changes in the backbone. On the other hand, these changes rarely occur in regular secondary structures such as alpha helices and beta sheets. Thus, it is more appropriate to "shift" all insertions and deletions to turn and loop structures, which can solve the problem of insertions and deletions; however, changes in loop conformation are difficult

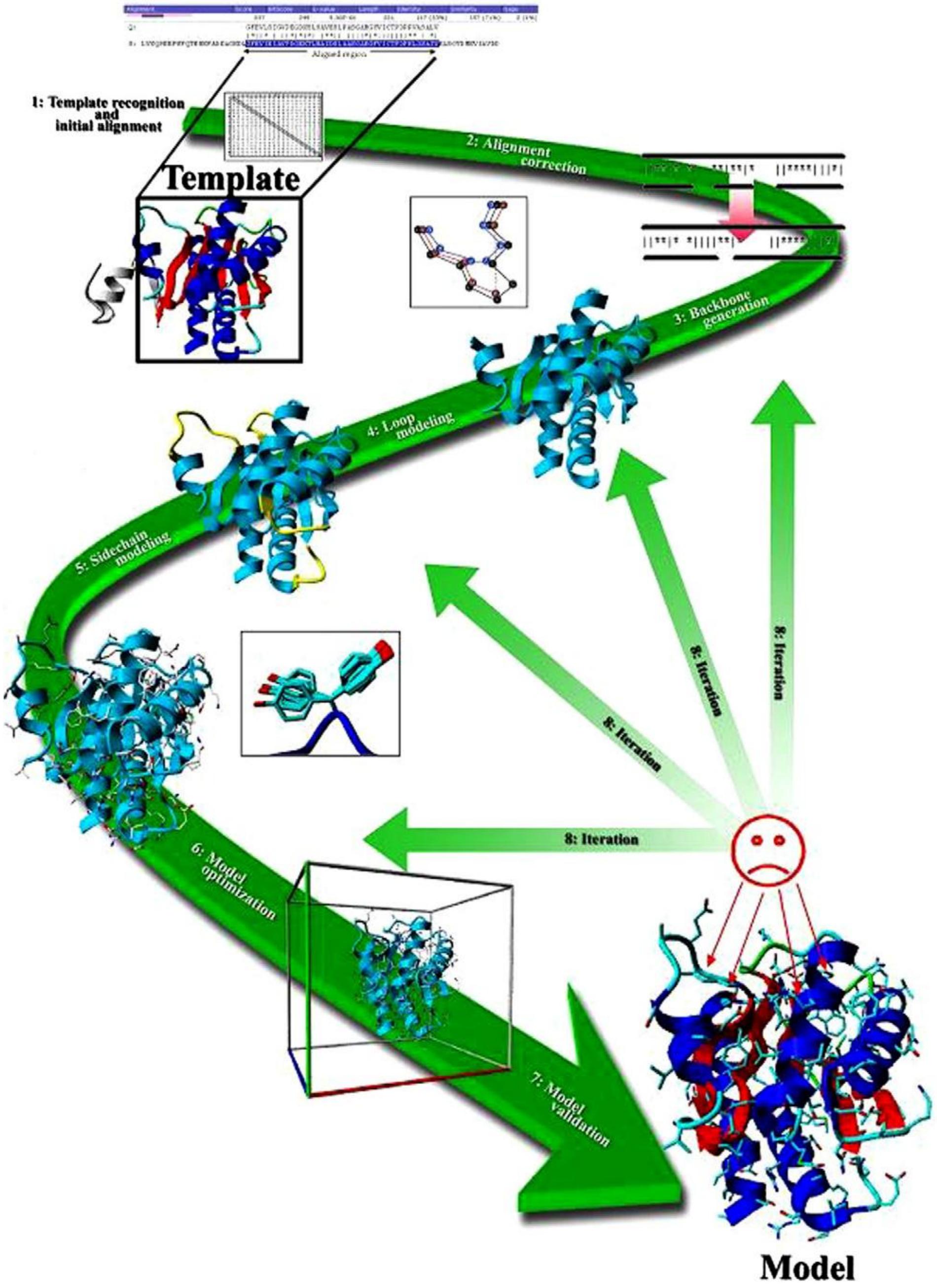

Aligned region
1: Template recognition and initial alignment
Template
2: Alignment correction
3: Backbone generation
4: Loop modeling
5: Sidechain modeling
6: Model optimization
7: Model validation
8: Iteration
8: Iteration
8: Iteration
8: Iteration
Model

FIGURE 13.9 **Homology modeling.** *Source: Based on illustration in Gu J, Bourne PE. Structural Bioinformatics, 2nd ed. Wiley-Blackwell, 2009. Used with permission.*

to predict. There are two main approaches to modeling loops: knowledge based and energy based.

Knowledge-based loop modeling searches the Protein Data Bank repository for loops of the same length and similar end-point geometry. Later the coordinates of the identified loop are transferred. Energy-based loop modeling utilizes an energy function to determine the quality of the loop and find the best loop conformation; it then performs a mineralization of the structure using either Monte Carlo [40] or molecular dynamics [41] techniques.

13.4.4.5 Side-Chain Modeling

In homologous proteins, where the percentage sequence identity is over 40%, the side-chain conformations (rotamers) of conserved residues exhibit the same orientation [42]. Therefore, it is possible to directly transfer the conserved residues from the template sequence to the model sequence, as mentioned earlier. However, this is only applicable with high levels of sequence identity. To overcome this problem, side-chain modeling uses a knowledge-based approach, where libraries of common side-chain conformations are employed to position side chains. Each conformation in the library is scored with a variety of energy functions and the best one is used. For this reason, in cases of low levels of similarity rotamer prediction can be computationally expensive.

13.4.4.6 *Model Optimization*

To predict side-chain conformation with high accuracy, one needs the correct main-chain conformations, which in turn depend on side-chain conformation. Thus, there is a loop in modeling main-chain and side-chain conformations, a problem solved by iteratively modeling the structure of the main chain and the side chain until the solution converges: First, the structure of the rotamers are predicted; following that, the structure of the main chain is remodeled to accommodate rotamers; finally, the rotamers are refitted to a new main-chain structure in a process that consists of a series of rotamer predictions and energy minimization steps. Another approach to optimizing the model is simply running a molecular dynamic simulation of it. This simulation follows the motions of the protein and mimics the folding process itself.

13.4.4.7 *Model Validation*

Errors in homology modeling depend on two factors: the percentage of sequence identity between the template and model sequences and the number of errors in the template sequence. If the identity between template and model sequences is higher than 90%, the accuracy of the predicted model is comparable to experimentally determined structures [32]. An identity percentage of 50–90% leads to a model with considerable local error, while an identity percentage of 25–50% often leads to large error regions, in which case homology modeling is not applicable.

Errors in the protein structure can be estimated either by computing the model's energy or by determining the normality indices. The former checks the bond length and the bond angles and determines if they are within the normal range, while the latter compares the characteristics of the predicted structure with those of the real structure. These characteristics include distribution of polar and nonpolar residues, direction of atomic contacts, and so forth.

13.4.5 Fold Recognition

Homology modeling is useful when the model protein (with a known sequence and an unknown structure) is related to at least one other protein with both a known sequence and a known structure. The quality of the predicted structure by homology modeling depends on the degree of similarity between the model and template sequences. If the similarity is very low, homology modeling of the query protein does not yield a meaningful result. Fold recognition can be employed in this situation. Its goal is to answer this question: Given a library of known structures and a model sequence, does the query protein share a fold with any of the known structures in it? One of the fold recognition methods is called threading and is discussed in the next section.

13.4.6 Threading

The idea behind threading is simple: based on each known structure in the library, build a rough model of the query protein and explore different possible sequence alignments for it and for the known-structure protein. Threading thus systematically explores many different alignments to find the best model. Both threading and homology modeling utilize sequence alignment of the query protein and the known-structure protein to predict the former's 3D structure. Table 13.3 compares homology modeling and threading.

Since threading builds one model per each known structure in the library, selecting the best requires scoring. Moreover, a way to evaluate the scores is required; otherwise, it is not clear whether a scoring model is correct.

One effective scoring approach is derived from patterns of residue neighbors in the library of known structures. First, the distribution of inter-residue distances for all 20 × 20 residues is retrieved from the known structures. Next, for each pair of residues, a probability distribution is derived based on the separation of the pair in physical space and in the amino acid sequence. Once the probability distribution for each pair is computed, one can estimate the performance of different models based on the level of agreement between them and known-structure distributions. The Boltzmann equation, which relates probabilities and energies, can be employed to derive an energy function from the probability distribution. This energy function can be used to evaluate the score of the threading model.

TABLE 13.3 Comparison of Homology Modeling and Threading

Homology Modeling	Threading
Identify homologous proteins	If no homologous proteins exist; try all present folds in library
Determine best alignment	Explore various possible alignments
Optimize only one model	Find best model among many rough models

Note: Homology modeling deals with one set of alignments; threading, many possible alignments. The goal of homology modeling is a very detailed model; threading produces a rough model.

13.4.7 *De Novo* Structure Prediction

The goal of *de novo* structure prediction is to predict the 3D structure of the query protein only from its amino acid sequence. In contrast to homology modeling and fold recognition, *de novo* prediction does not rely on homology between query sequence and known-structure sequence. Thus, it considers a noticeably larger conformational landscape than homology modeling and fold recognition, since both of latter only explore regions of space near an initial template structure. Sampling this large landscape is computationally expensive, to the extent that it can be considered a limitation of de novo methods. Global search algorithms, such as Monte Carlo Metropolis, can be employed to ensure that important regions of the conformational landscape are sampled. *De novo* methods are also known as *ab initio* methods.

De novo methods have been shown to be successful for predicting secondary structures and local structure patterns [43,44], although, despite recent progress, there are many problems that need to be addressed before a consistent and reliable prediction can be obtained.

13.5 BIOINFORMATICS DATABASES

In the post-genome era, there is not only a wide range of database content but also a vast amount of choice within each discipline or data type—from extremely large and broadly themed international repositories to databases solely devoted to, for example, the targets of an individual transcription factor [45]. It is advantageous for researchers in most disciplines in medicine and research biology to exploit at least some of these valuable resources.

The aim of this section is to (1) identify key fields where large databases are influential in the biosciences; (2) very briefly explain why they are useful and describe their scope; and (3) outline some of the most prominent and useful databases. The two most valuable resources are the National Center for Biotechnology Information (NCBI) which is part of the National Institutes of Health (NIH), and the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI). Much of the information given here is provided by these organizations, and both provide the option of a general search of many of their databases through one interface (Entrez, Ensembl). File types are varied throughout IT and bioinformatics, and some are individual to just one database. The most common will be outlined, with the exception of pdbs, which were described in Section 13.4.2.

13.5.1 General Sequence Databases

Compared to older technologies, modern massively parallel sequencing technologies are capable of generating a vast amount of information at a fraction of the cost, and they have revolutionized genetic research. This explosive growth in fact exceeds the growth of storage capacity and has resulted in new challenges for data storage, retrieval, and analysis [46]. The primary resource for raw sequence data, both DNA and RNA, is the Sequence Read Archive (SRA), a public repository of data from next-generation sequencing (NGS) machines.

The SRA is part of the International Nucleotide Sequence Database Collaboration (INSDC), which is a collaboration of three partners: EBI, NCBI, and the DNA Data Bank of Japan (DDBJ) [47]. Each partner uses its own graphical user interfaces to provide electronic access to the database stored on their servers. A programmatic way to access the information is through the R/Bioconductor package, SRAdb [48], which allows the user to search through an SQL database of SRA metadata, download the appropriate data, and interactively view it through the Integrated Genomics Viewer (IGV). The metadata include information about the study, sample, experiment, run, analysis, and submission. EBI and NCBI also offer programmatic methods of accessing their database (using the ENA browser REST URLs and the SRA System Development Kit, respectively); however, Bioconductor contains many useful analysis packages and so it may be more convenient to use SRAdb.

The SRA accepts base calls, quality scores, and sequence alignments through any of its three partners and, in 2011, had accumulated over 100 terabases of "reads" (a term often used in concordance with NGS technologies to denote a short contiguously sequenced length of bases). Because of the number of currently available technologies, the number of upcoming technologies, and the variety of associated file formats, compatibility is an issue. The solution is that data in some (newer) file formats are provisionally available only as they are, whereas data submitted in already widely used formats is validated and made available in standard formats. Regardless of which of the three SRA partners a dataset is uploaded to, it is shared among all. As a result, each database is almost identical, although methods of access and databases linked to the SRA data vary.

The largest individual project in the SRA is 1000 Genomes, which is an international effort and contributes almost a third of the bases in SRA [49]. The aim of the project is to investigate human genetic variation across geography and phenotype, and in so doing promote understanding of the genetic contribution to human disease. The 1000 Genomes Project Consortium describes the genomes of 1092 individuals from 14 populations throughout the world. The pilot phase—in which >95% of common variation (>5% frequency) was discovered—showed that a combination of methods could ensure a balance between cost-effectiveness and power to describe all but the rarest of single-nucleotide polymorphism (SNP) and indel (insertion/deletion) variants. The methods employed were low-coverage whole-genome sequencing, deep exome sequencing, and extensive SNP genotyping. The consortium reported 1.4 million short indels, more than 14,000 large deletions and 38 million SNPs, and covered up to 98% of accessible SNPs at a frequency of 1% in related populations.

It can be said then that 1000 Genomes provides extensive coverage of common variants and high coverage of low-frequency variants, but poor coverage of the rarest variants. As well as providing a resource for molecular research in general, the consortium has uncovered interesting summary information about the genome—for example, that each individual "carries" hundreds of rare noncoding variants at conserved sites which potentially alter transcription factor binding or have other expression-related effects. While 1000 Genomes data is part of the SRA, it also has a stand-alone website (www.1000genomes.org). In addition, it submits its findings to genetic variation databases such as NCBI's dbSNP (www.ncbi.nlm.nih.gov/SNP).

EBI makes SRA accessible through the European Nucleotide Archive (ENA) [50]. In 2013, ENA published updates to the effect that the SRA had tripled in size since 2011 and was consistently doubling in size every ten months; it now comprised data from over 14,000 studies. Furthermore, comparable growth is expected for another five to ten years, even under modest predictions [51]. In addition to the SRA, ENA also has the Trace Archive and the EMBL-Bank. The Trace Archive is similar to SRA: it is a large repository of raw nucleotide sequences maintained by each INSDC partner; however, it is used primarily for older lower-throughput gel/capillary sequencing, which typically results in longer reads but lower overall throughput than those obtained by NGS technologies. EMBL-Bank is the continuation of the EMBL Nucleotide Sequence database, containing nucleotide sequences and their associated biological annotations and bibliographic information. Unlike the SRA or Trace Archive data, nucleotide sequences housed in EBML-Bank are processed rather than raw—although many of the data are based on the raw data in the SRA and the Trace Archive.

NCBI's GenBank is similar to EMBL-Bank as a repository for nucleotide sequences and their corresponding biological and bibliographic annotations. Although these two resources are not strictly part of the INSDC, they do share data with each other and with DDBJ on a daily basis [52]. GenBank itself subdivides into three smaller databases: Expressed Sequence Tags (EST), Genome Survey Sequences (GSS), and the Nucleotide database, which holds all the remaining data, including whole-genome shotgun data, representing most of the data in GenBank.

When using any of the aforementioned databases, it should be remembered that entries are submitted by thousands of researchers internationally and are not extensively reviewed. There is, therefore, both an extremely high level of redundancy and a varying level of quality and correctness of the associated annotations. Another useful sequence database maintained by NCBI is RefSeq [53], a nonredundant database that is highly curated and, while not part of the INSDC, derives nonredundant sequences from INSDC data. Similarly UniGene, rather than describing genes, as the name implies, nonredundantly describes all known transcripts that originate from a particular locus.

All of these resources are described as open access; that is, access is free and unrestricted. In some experiments, where consent documents have not included open access, data sharing must be restricted. For example, a restriction may be that the data are only to be used for specific research or provided only to bona fide researchers. In these cases, patient confidentiality must of course be maintained with appropriate security measures. To avoid mistakes, access should be granted only by the access-granting organization (DAO) and not by the database administrators. EBI's European Genome-Phenome Archive (EGA) and NCBI's database of GenoTypes and Phenotypes (dbGaP) are examples of such databases; they store data on entire experiments, including sensitive but well-characterized patient phenotype information corresponding to the experimental data [54].

Datasets from these resources can potentially be used for a range of applications, from testing new or developing software on appropriate real data to conducting meta-analysis over several datasets of the same type. There are pitfalls, however. One problem is that limited public reporting of secondary research on these datasets to some extent obscure the true extent of its impact. Another is that no consensus exists as to whether additional consent is required to submit the datasets to such a database, even with security restrictions. In a survey about this issue, participants stated that it was very (69%) or somewhat (21%) important that they be asked permission, indicating ethical consent should be a concern [55].

13.5.2 Gene Expression

One of the many technologies heavily influenced by the advances of the Human Genome Project is the microarray. Microarrays are flat surfaces or "chips" on which a 2D array of probes is located, all of which are used to assay the same sample. In some cases, there are several sets of probes per microarray and many samples may be assayed. Each "spot" on the array tests for the presence of a specific biomolecule of a particular type, usually DNA. The probes, which fluoresce when binding their target molecule, are hybridized either to the surface (as in Affymetrix technology) or to microscopic beads located in wells on the surface to help resolve problems arising from spatial technical effects (as in Illumina technology). These devices are perhaps best known for assaying gene expression, although they have many other uses, including the investigation of protein binding and DNA methylation. In practical terms, a gene expression microarray allows the researcher to see which genes have an altered expression—for example, depending on different phenotypes or drug doses.

NCBI's GEO [56] and EBI's ArrayExpress [57] are the main databases for array-based data, including various types of microarray and some sequence-based data such as RNA-seq. RNA-seq is essentially a method for estimating gene expression levels by sequencing cDNAs and aligning them. It is generally seen as a more powerful tool, and it has been suggested that it will replace gene expression microarrays as the cost of sequencing decreases [58]. Microarrays are still a valuable technology, being cheaper, less complicated, less data-heavy, and simpler to analyze. RNA-seq, however, can detect SNPs and other genetic variations, fusion genes, and levels of different splice variants. It has also been suggested that the two technologies should be used together where possible to detect the maximum coverage of the transcriptome [59]. All data in GEO and ArrayExpress must adhere to "minimum information about a microarray experiments" (MIAME) [60] or "minimum information about sequencing experiments" (MINSEQE) [61]; these are standards dictating that enough information must be included for a second party to sufficiently analyze and/or recreate the experiment. In 2012, ArrayExpress reported having over 30,000 datasets—27,000 more than it had in the previous two years. Both GEO and ArrayExpress continue to grow quickly, and both tools, along with ways to link them to popular tools, are in development [62].

In some cases, rather than interrogating a specific case-control experiment dataset, one might simply wish to know which tissues normally express a gene and to what extent. BioGPS allows searching by several types of identifiers, including gene symbol, gene ontology, Affymetrix ID, and genomic interval [63]. The result is a bar chart of expression values for a list of tissues that have been assayed by Affymetrix microarrays. Reports of various annotations and a gene summary are also shown alongside the data. To identify genes that are normally similarly expressed and so potentially share similar roles or regulatory pressures, their expression profile may be searched against other expression profiles for one with a correlation surpassing a preselected cut-off. Another, more comprehensive take on the theme, is EBI's Gene Expression Atlas [64]. This web service offers similar searches and output, but across a variety of biological conditions and diseases as well as across tissue and cell types. Graphical displays are employed, and the user can obtain a view of many types of biological contrast from a single report page; the user can also see, for each tissue, the number of studies in which the gene showed up- and down-regulation, and obtain links to the data in ArrayExpress as well as in publications describing the experiments.

13.5.3 Noncoding Sequences and Epigenomics

The term "junk DNA" dates back to 1972 and describes the large proportion of noncoding DNA in the human genome which at the time had no known function [65]. Since then, many types of regulatory noncoding genetic sequences have been recognized. Some of this "junk DNA" is in the form of repetitive elements that may span one-half to two-thirds of the genome [66]. Repetitive DNA has been shown to perform a variety of functions, one of which is serving as a boundary for heterochromatin domains [67]. It has also been shown to influence various diseases—for example, Huntington's disease, which is caused by a trinucleotide repeat in the gene HTT [68]. Repbase Update is the most commonly used database for repetitive DNA elements; it reports on a diverse range of repeat families and subfamilies [69] and contains over 3600 annotated repeat sequences, some unique to it. Repbase Update may be used either to identify genomic locations of repeats or to screen them from a sequence so that only the nonrepetitive regions are left uncensored (repeat bases are converted to the letter "N").

Promoters are genetic sequences that are usually located upstream of their corresponding genes; they are usually a few hundred base pairs in length and are involved in the initiation of transcription, where they act as a binding site for the transcription initiation complex. Some genes have multiple promoters, and different promoters for the same gene may be activated under different circumstances, causing differential expression [70]. The Eukaryotic Promoter Database (EPD) contains experimentally defined promoters of RNA polymerase II (the enzyme responsible for mRNA synthesis) [71]. These data have been collected over 25 years, beginning with experiments mapping transcription start sites (TSS) for individual genes and later the use of carefully selected NGS datasets. The focus of EPD was originally on high quality, whereas a new section, called EPDnew, focuses more on quantity, with the aim of identifying promoters for all genes over several model organisms [72]. Another recent addition is the use of chromatin signatures and mRNA 5′ tags to help identify genes with lower expression.

Transcription factors (TFs) are proteins that bind to DNA, thereby influencing gene expression. Those that promote the recruitment of RNA polymerase up-regulate a gene and are said to act as activators; those that block recruitment and down-regulate are said to be repressors. Possibly the most popular database for TFs is Transfac; however, only an older database version with an interface of reduced functionality is available to the public [73]. Transfac provides data for experimentally proven TFs and their corresponding genes, and provides predictions for potential TF-binding sites. One recently published alternative database is AnimalTFDB, which allows browsing by TF family and species as well as the use of free search terms [74].

MicroRNAs (miRNAs) are short noncoding RNA molecules (~22 nucleotides in length) that function in both transcriptional and post-transcriptional eukaryote regulation [75]. They are known to "silence" or "knock down" genes either by degrading mRNA or by preventing transcription or translation via binding to the corresponding DNA or mRNA, respectively [76]. MirBase is probably the most useful place to initially look for miRNA annotations, sequences, and so forth. Each entry has a thorough report page including sequence, genomic coordinates, and links to scientific papers on PubMed. Unfortunately, MirBase does not provide information on target genes [77], but many websites provide this functionality, describing experimentally validated targets, targets predicted by various algorithms, or both. A recent web application, miRWalk, summarizes the information from eight of these databases [78].

There are many definitions for epigenetics, a recent one being "stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" [79]. Epigenomics is the study of the complete set of at least one of these changes and, functionally speaking, epigenetics experiments often use chip-based or sequencing-based technologies to investigate DNA methylation or histone modification. The Epigenomics database at NCBI stores data for over 4000 experiments focusing on various species, biological samples, and "feature types" (i.e., epigenetic modifications) [80]. It is essentially a subset of GEO, but with a modified interface specifically for those searching for epigenomics data.

13.5.4 Proteins, Metabolites, and Pathways

The Universal Protein Resource (UniProt) is a comprehensive database of structure and annotation and is divided into four main components: UniProt Knowledgebase (UniProtKB), UniProt Reference Clusters (UniRef), UniProt Archive (UniParc), and UniProt Metagenomic and Environmental Sequences (UniMES) [81]. UniProtKB is divided into a TrEMBL section, which stores automatically annotated and unreviewed sequences, and a Swiss-Prot section, which stores manually annotated and reviewed sequences. UniProtKB/TrEMBL contains ~40,000,000 sequences comprising ~13,000,000,000 amino acids; UniProtKB/Swiss-Prot contains ~540,000 sequences comprising ~190,000,000 amino acids.

UniProtKB/Swiss-Prot is less redundant in that it stores only one entry and ID corresponding to each gene. UniProtKB/TrEMBL may store several entries per gene depending on nonsynonymous genetic variation, alternative splicing, and fragmentation (these entries are simply the translations of sequences found in EMBL-Bank/GenBank/DDBJ). For this reason, Swiss-Prot has become a valuable tool for proteomics, where very minor variations are not highly important; often more important is that corresponding proteins be given common names to simplify analysis and literature/database searches and that protein levels be compared (e.g., between case and control groups). Data from over 28,000 proteomics experiments can be found and downloaded from the Proteomics Identification Database (PRIDE), and UniProt can be used to convert redundant IDs into SwissProt IDs [82]. The TrEMBL section, on the other hand, may be more useful for investigating differences in structure and function due to nonsynonymous genetic variation/mutation/alternative splicing. Each report page offers a wealth of information on the protein of interest, including post-translational modifications and links to other resources, such as the Protein Data Bank [83].

UniParc is an archive of protein sequences extracted from public databases; however, no report pages are given and all associated data must be retrieved through links to the various sources. UniRef stores clustered sets of sequences from UniProtKB and UniParc in such a way that searches result in clusters of related proteins (e.g., isoforms and proteins from orthologs). Unlike UniParc or UniProtKB/TrEMBL, UniRef merges sequence fragments to reduce redundancy and simplify interpretation; however, isoforms are kept as separate records to include more "resolution" than is available in UniProtKB/SwissProt. In a sense, one can say that none of the UniProt databases are redundant and that each uses its own definition of redundancy—redundant by gene, redundant by splice variant, and so forth. The UniMes database contains clusters of sequences from metagenomics clusters obtained from experiments such as Craig Venter's Global Ocean Sampling Expedition; the sequences cannot be found anywhere in UniProt except in UniParc.

The Human Metabolome Database describes small molecule metabolites in humans, containing and linking three types of data: chemical, clinical, and molecular biology/biochemistry [84]. It holds more than 40,000 metabolites, both water-soluble and lipid-soluble. Some of them are called "detected" because they have been experimentally confirmed, while others are called "expected" because of their biochemical pathways or high human exposure. It is possible to browse by biofluid or pathway, or to search via a range of methods, from peak matching to chemical formula. The standard report page for a particular metabolite gives its description, chemical taxonomy, biological and physical properties, and links to various useful sites such as Kyoto Encyclopedia of Genes and Genomes (KEGG) and NCBI's PubChem (pubchem.ncbi.nlm.nih.gov/).

Various interaction and pathway databases can be important for downstream analysis of a wide range of experiments. Some are simply for storage and searching whereas others provide extra functionality. KEGG Pathway is a large well-curated database with impressive zoomable "global" maps that show large sections of metabolism in one view. However, part of its service is now commercial [85]. A similar database with fewer compounds but more pathways and reactions, and slightly more extensive pathway report pages, is MetaCyc [86]. Pathway Commons collects data from a range of free-access databases and provides a link for interactive pathway viewing by Cytoscape [87]. Reactome allows mapping of a list of IDs to its database of pathways, allowing the user to see if a set of biomolecules of interest are heavily involved in one pathway in particular [88]. PathVisio is a tool for viewing and editing the pathways in WikiPathways [89,90], which is a noncommercial site unrelated to Wikipedia but is also an open and collaborative platform. Because PathVisio is written in Java, it is cross-platform; also, its website provides user tutorials.

The pathways just discussed are only a few of the choices available, and only some of their "front-end" functionality was described. To give the reader an idea of the scope of available resources, it is worth mentioning Pathguide (http://pathguide.org), which describes itself as a "pathway research list." A search of their website with no terms selected to narrow it (e.g., by organism) yields "556 biological pathway related resources and molecular interaction related resources."

13.5.5 Pharmacogenomics

As NGS costs drop and more genetic variation becomes annotated, personalized medicine becomes more realizable. With this in mind, Pharmacogenomics Knowledge Base (PharmGKB) has been set up to collect information about genetic variation and its influence on drug response, and to make this information freely available [91]. PharmGKB makes use of a range of data, from genetic to clinical, and the user can browse it by gene, drug, disease, or pathway; each gene, drug, or disease record is marked to indicate how fully annotated it is—for example, whether there is clinical data or whether genetic variants are annotated as being associated with a particular drug. Over 100 pathways are curated by PharmGKB (in addition to links) and have thorough descriptions that include external links to evidence and various options for downloads and internal links for pathway components. There is also a search bar with a Google-like auto-complete function for finding drugs, genes, and so forth. PharmGKB and the NIH Pharmacogenomics Research Network (PGRN) are making a collaborative effort to develop genetics-based drug dosing guidelines. PharmGKB recently added these data to the database and correspondingly updated its web interface.

13.5.6 Annotation

One of the most useful annotations in biology is gene ontology (GO) [92]. GO terms represent a standardized classification of genes and their products and are split into three broad types: biological process, cellular component, and molecular function. Each term is part of a hierarchy—for example, "mRNA splicing, via spliceosome" is a child term of "mRNA processing." Gene ontologies are useful for term enrichment analysis where an algorithm detects if any terms are over-represented in a list of genes of interest, such as differentially expressed genes. The more similar a model, the more useful, which is why gene ontologies are not static but changing, with each version subject to the Gene Ontology Consortium's review and benefiting from the scientific community's advice. When a term is associated with a gene, an evidence code (e.g., "IMP: inferred by mutant phenotype" or "ISS: inferred from sequence or structural similarity") is also recorded so that the reader or algorithm may filter out associations with weaker evidence.

Online Mendelian Inheritance in Man (OMIM) is an extensive database of diseases with genetic components [93]. Its entries have MIM numbers—unique 6-digit identifiers wherein the first digit describes the method of inheritance (e.g., X-linked, autosomal). A character before the code describes additional details, and a number after a decimal point at the end of the code denotes an allelic variant.

The NCBI Taxonomy Database is for INSDC members and was developed in consultation with those members and external advisors [94]. Almost 280,000 species have been described by Taxonomy, which represents about 10% of the described species of life on the planet. Finally, GeneCards (www.genecards.org/) is a very useful source of summary information on a gene of interest, and it provides relevant links to many of the databases mentioned here.

13.5.7 Common File Formats

There are many standard file formats found throughout informatics and many custom file formats unique to bioinformatics. The most simple is delimiter separated values (DSV), in which rows are separated by an end-of-line (EOL) character (generated by the return key). There are system-specific (Windows, Mac, Linux) ways for encoding these characters, although the software often recognizes any method, and its conversion is easy with a good text editor such as Notepad++ or Sublime. Columns are specified by a so-called *delimiting* character. For example, in tab separated value (TSV) files, tabs separate columns; in comma separated value files, CSVs commas are used. One of the strengths of this format is that it can be opened in MS Excel as a spreadsheet or can be relatively easily read and edited in a text file. Another strength is that the delimiter is easily changed and generally may be imported into many programs. Weaknesses include the situation where the data contain a delimiter, effectively splitting the containing cell in two. This can be overcome by surrounding each entry with quotes so that every delimiter character in the data is within quotation marks and so is interpreted correctly, i.e. without special meaning. Also, the DSV format is without any internal structure beyond that of its rows and columns. Sometimes an additional delimiter is used to allow a varying number of one attribute to be linked to another; for example, the gene column of a particular row might contain `“CHID1”` and the probes column might contain `“gnf1h00419_at| gnf1h00420_s_at| gnf1h09948_x_at”`, where a vertical bar “|” separates the data. This can be useful with simple relationships, but it is not commonly supported and, for datasets with more complex relationships it is certainly advisable to use Extensible Markup Language (XML) instead.

XML documents open with an XML declaration, which provides the XML version and character encoding used. The declaration of character encoding indicates which character set was used so it can be correctly read by the recipient. The old web standard—that is, American Standard Code for Information Interchange (ASCII)—is still widely used despite its limited range. To include characters outside of ASCII, such as Greek letters, a Unicode character set, such as UCS Transformation Format—8-bit (UTF-8), is required; it uses multiple bytes to represent additional characters. In XML a parent element may have any number of child elements, which may themselves have any number of child elements and so on. In this case, every entry is surrounded by or ‘wrapped in’ case-sensitive tags. An XML tag is a string of characters itself wrapped in ‘lesser than’ and ‘greater than’ signs (i.e., < >), which respectively indicate the beginning and end of the tag. The second tag also has a forward slash immediately after the “<” to indicate that it is a ‘closing tag’ rather than an ‘opening tag’. This is similar to Hypertext Markup Language (HTML) in structure but not in purpose; moreover, XML tags are defined by the file’s creator rather than by a universal standard. Tags may also have associated attributes that act as classifiers. White space (i.e., spaces, tabs, new lines) outside the tags is meaningless and so can be helpful in formatting the file for easy viewing; this often indicates the nested relationships of the tags. Many good text editors, such as Notepad++ (http://notepad-plus-plus.org) or Sublime (www.sublimetext.com), will automatically color-code files to make them more readable when viewing directly as text. For example:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!-- This is a comment, e.g., “Last edited 24/07/2013 – SWR” -->
<genome species=”human”>
        <gene>
                <name>Chitinase Domain Containing 1</name>
                <ID>CHID1</ID>
                <probe type=”Affy”>gnf1h00419_at</probe>
                <probe type=”Affy”>gnf1h00420_s_at</probe>
                <probe type=”Affy”>gnf1h09948_x_at</probe>
                <probe type=”Illumina”>ILMN_2069615</probe>
        </gene>
        <gene>
                <name>NAD Kinase</name>
                <ID>NADK</ID>
        </gene>
</genome>
```

For more complex relationships, there are relational databases. These are collections of tables that follow the relational model [95]. They must be accessed through a relational database management system (RDBMS) such as MySQL. Relational databases comprise a number of tables, but can often be saved into a single file—such as MySQL’s dump file—by their RDBMS and restored on another machine.

Each table must have a primary key: a column with data that uniquely identify each row, or a group of columns whose combination accomplishes this. The latter is known as a compound key because it is made up of several columns. A foreign key, on the other hand, allows records in two tables to be related to one another by a relationship. A 1–n (one-to-many) relationship, as seen in the `gene` and `probe` elements in the XML example, may be described in such a manner. The power of a relational database comes from its ability to support complex m–n (many-to-many) relationships using a third table. As well as data storage in a structure representative of the real world it describes, an RDBMS often has built-in security and Structured Query Language (SQL), subtle variants of which are used by different RDBMSs. Simple SQL queries select certain cells of an attribute from a table where another attribute in that row is equal to a value of interest. More complex queries involve the use of relationships between tables to appropriately join records. The following is a typical, simple example:

```
SELECT name FROM gene WHERE ID=“CHID1”
```

All of the aforementioned file types are common throughout IT. FASTA files are specific to nucleotide and protein sequences and describe either one or multiple sequences. A description line indicated by a “>” serves as a sequence identifier. The sequence follows its corresponding description line. NCBI recommends that each line be restricted to a maximum of 80 characters. Numbers on the left of each column may be present to indicate the position of the first base of that row relative to the first base in the sequence. Spaces can be used to more easily identify the position of each base in the row (e.g., 1 space every 10 bases). File formats such as FASTA are more easily viewed using a monospaced font, where each character occupies the same width across the page such that the columns of letters are aligned. Monospace fonts may also be useful for comparing SQL queries and looking for spelling mistakes that result in errors. The following is an example from the NCBI FASTA specification (www.ncbi.nlm.nih.gov/BLAST/blastcgihelp.shtml):

```
>gi|129295|sp|P01013|OVAX_CHICK GENE X PROTEIN (OVALBUMIN-RELATED)
qikdllvsss tdldttlvlv naiyfkgmwk tafnaedtre mpfhvtkqes kpvqmmcmnn
sfnvatlpae kmkilelpfa sgdlsmlvll pdevsdleri ektinfeklt ewtnpntmek
rrvkvylpqm kieekynlts vlmalgmtdl fipsanltgi ssaeslkisq avhgafmels
edgiemagst gviedikhsp eseqfradhp flflikhnpt ntivyfgryw sp
```

The FASTQ format was developed at the Sanger Institute at the turn of the century and has since become a standard for nucleotide sequencing [96]. There are four line types in FASTQ. The first is the title, which is similar to the description line in FASTA and is indicated by "@" followed by an identifier for the read. Next comes the line for the sequence itself, which, when viewed in a text editor, may spill over if the editor uses a fixed width; however, there should be no EOL characters. The sequence is usually written in uppercase and in standard characters unbroken by white space. The third line begins with "+" and can optionally be followed by the same identifier as in the title line for clarity. The final line contains a length of characters, identical to that in the second line, with each character indicating the quality of the corresponding nucleotide on the second line. There are three FASTQ variants each with different quality scores or encoding of the scores but all of them rely on nonnumeric characters to represent numbers. For example, the original (Sanger) Phred score coding ranged from 0 to 93 and was encoded with ASCII codes 33 (!) through 126 (~)

```
@SRR014849.1 EIXKN4201CFU84 length=93
GGGGGGGGGGGGGGGGCTTTTTTTGTTTGGAACCGAAAGGGTTTTGAATTTCAAACCCTTTTCGGTTTCCAACCTTCCAAAGCAATGCCAATA
+SRR014849.1 EIXKN4201CFU84 length=93
3+&$#"""""""""""7F@71,’";C?,B;?6B;:EA1EA1EA5’9B:?:#9EA0D@2EA5’:>5?:%A;A8A;?9B;D@/=<?7=9<2A8==
```

This range includes upper- and lowercase letters, numbers, and punctuation and excludes white space. One issue with this arrangement is that the "@" and "+" characters are included in the range. This is why it is beneficial to exclude additional EOL characters from FASTQ files to reduce the necessary complexity and the likelihood of errors made by parsing algorithms. Phred is defined as

$$Q_{\mathrm{PHRED}} = -10 \times \log_{10}(P_e)$$

Here "P_e" represents the probability of an erroneous base call. Phred scoring was introduced by the Phred software program, which reads DNA sequencing trace files, "calls" bases, and assigns quality scores. Phred is also used for Illumina FASTQ files from Genome Analyzer Pipeline version 1.3 onward, but with a different range and a different set of ASCII values, so it is easily converted. Software is available to easily convert between these two scoring systems and the scoring system of the third non-Phred FASTQ variant ("Solexa"—the early Illumina score).

GenBank, EMBL, and DDBJ sequence formats follow a similar style and share the taxonomy of NCBI's Taxonomy project and the Feature Table section of each file type for which the formatting is also standardized across the INSDC (www.insdc.org/files/feature_table.html). These formats include annotation of the sequence and its corresponding organism, and they are convenient both for visual inspection in a text editor and for parsing by a computer program. They may describe, for example, a gene or an entire bacterial genome. The EMBL sequence format uses a two-letter code followed by three spaces to begin each line except the sequence lines themselves. The first line always uses the "ID" code for identification of the entry (e.g., accession number, sequence version number, topology.) and no other ID codes for that entry; however, many following line types can be ignored or used several times. The first set of lines annotates the sequence, the organism the sequence is from, pubmed references to relevant papers, and so forth.

A Feature Table follows, identified by `FT` and describing regions or sites of interest. For example, genes and their exons and coding sequences might be notated, using base positions for start and end points for genetic sequences and translated sequences for the corresponding proteins. At the end of the file, the `SQ` (sequence header) line code denotes the header for the sequence. The sequence itself follows, with 60 bases per line and split every ten bases by a space for readability, and ends with `//`. RNA sequences are stored as they exist *in vivo*; that is to say, it is the RNA code, not the corresponding DNA code for the cDNA that was sequenced. The file may be separated by blank "`XX`" lines simply for readability and "`CC`" lines to allow free text comments. The GenBank and DDBJ formats are very similar, but are more self-explanatory because they use entire words such as "`LOCUS`," "`ACCESSION`," and "`FEATURES`" rather than codes and still meet the requirements for parsing in that they are standardized. Additionally, they contain

TABLE 13.4 Summary of Databases and Related Resources

Name	URL	Main Features
Entrez	http://www.ncbi.nlm.nih.gov/gquery	Searches NCBI databases
Ensembl	http://www.ensembl.org/index.html	Searches EBI databases
SRA	http://www.ncbi.nlm.nih.gov/sra	NGS read data
GenBank	http://www.ncbi.nlm.nih.gov/genbank/	Annotated sequences
Trace Archives	http://www.ncbi.nlm.nih.gov/Traces/home/	Gel/capillary sequencing data
ENA	http://www.ebi.ac.uk/ena/	The EBI equivalent of the three above combined
RefSeq	http://www.ncbi.nlm.nih.gov/refseq/	Non-redundant, well annotated sequence database
EGA	https://www.ebi.ac.uk/ega/	Experimental datasets, some access restriction
dpGaP	http://www.ncbi.nlm.nih.gov/gap	Experimental datasets, some access restriction
DDBJ	http://www.ddbj.nig.ac.jp/	Third INSDC partner
dbSNP	http://www.ncbi.nlm.nih.gov/SNP/	SNP data
GEO	http://www.ncbi.nlm.nih.gov/geo/	Large functional genomics data bank
ArrayExpress	http://www.ebi.ac.uk/arrayexpress/	Large functional genomics data bank
BioGPS	http://biogps.org/	Expression profiles across tissues and states
Gene Expression Atlas	http://www.ebi.ac.uk/gxa/	Expression profiles across tissues and states
RepBase Update	http://www.girinst.org/repbase/	Repetitive DNA
EPD	http://epd.vital-it.ch/	Promoter database
Transfac	http://www.gene-regulation.com/pub/databases.html	Transcription factor database
AnimalTFDB	http://www.bioguo.org/AnimalTFDB/	Transcription factor database
miRBase	http://www.mirbase.org/	miRNA general resource
miRWalk	http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/	miRNA gene targets
Epigenomics	http://www.ncbi.nlm.nih.gov/epigenomics	Epigenomics experimental data
Uniprot	http://www.uniprot.org/	Protein annotations and sequences
PRIDE	http://www.ebi.ac.uk/pride/	Proteomics experimental data
Human Metabolome Database	http://www.hmdb.ca/	Metabolomic experimental data
KEGG Pathway	http://www.genome.jp/kegg/pathway.html	Pathway database
MetaCyc	http://metacyc.org/	Pathway database
Pathway Commons	http://www.pathwaycommons.org/pc/	Pathway database
Cytoscape	http://www.cytoscape.org/	Bioinformatics software
Reactome	http://www.reactome.org/	Pathway database and analysis tool
PathVisio	http://www.pathvisio.org/	Tool for interacting with WikiPathways
WikiPathways	http://www.wikipathways.org/	Pathway database
Pathguide	http://www.pathguide.org/	List of pathway databases

(*Continued*)

TABLE 13.4 Continued

Name	URL	Main Features
PharmGKB	www.pharmgkb.org	Pharmacogenomics knowledge base
Gene Ontology	http://www.geneontology.org/	Standardised gene annotation
OMIM	http://www.omim.org/	Database of diseases with genetic component
Taxonomy Database	http://www.ncbi.nlm.nih.gov/taxonomy	NCBI's taxonomy standardisation
GeneCards	http://www.genecards.org/	Gene summary information and links

a more obvious built in hierarchy that uses space-indented line descriptors. For example:

```
REFERENCE   2  (bases 1 to 756)
  AUTHORS   Kreft,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-1992) J. Kreft, Institut f. Mikrobiologie,
            Universitaet Wuerzburg, Biozentrum Am Hubland, 8700 Wuerzburg, FRG
```

Table 13.4 contains a listing of bioinformatics-related databases and relevant resources.

GLOSSARY

Alpha helix and beta sheet common secondary protein structures

Amino acid substitution matrix a matrix describing how common each type of amino acid change is, to facilitate scoring of similarity between protein sequences

BLAST a popular algorithm for sequence alignment

Bootstrapping the scoring of a tree or nodes of a tree by resampling and counting outcomes of each pseudoreplicate dataset

***cis* configuration** the two alpha carbon atoms of the connected amino acids on the same side of the peptide bond

Clade/monophyletic group a group in which every descendant of a node is present

Delimiter separated value a plain text file in which rows are separated by EOL characters and columns by a specified delimiter (e.g., tab or comma)

***De novo* structure prediction** the prediction of a protein's structure without the use of templates

Discrete method a method that considers differences in nucleotides directly

Distance method a method that considers only the number of differences between nodes

Dot-matrix a graph representing similarities between two sequences

Gap penalty a penalty on a sequence alignment score, made for the introduction of a gap

Gene ontology a standardized way of classifying genes and their products

Heuristic a "rule-of-thumb" method; not as accurate as, but faster than, an exact method

Homology relation by descent, inferred by similarity. Often used synonymously with 'similarity'

Homology modeling the use of similar protein sequences with known structures in protein modeling

Microarray a surface or "chip" on which a large 2D array of probes is located, used to assay a sample in a high throughput manner

MIM number an ID from OMIM specifying a disease with a genetic component

Molecular evolution the study of the evolution of DNA, RNA, or proteins

Multiple-sequence alignment an alignment of several sequences, often used in phylogenetics and protein modeling

Orthology homology where the sequences have deviated by speciation

Paralogy homology where the sequences have deviated by gene duplication

PDB file a file listing coordinates for the atoms of a protein

Phylogenetics the identification of evolutionary relationships between organisms or between genes

φ and ψ angles rotation angles of certain bonds of a protein backbone

Protein modeling the estimation of a protein's structure using its sequence

***trans* configuration** the two alpha carbon atoms of the connected amino acids are on the opposite sides of the peptide bond

Relational database a database with a complex structure, driven by a relational database management system (RDBMS)

Threading aligning a protein sequence to known structures

Time complexity describes the relationship between the input size n and the time taken for an algorithm to run

XML a text file with a hierarchical structure

REFERENCES

[1] Benton D. Bioinformatics—principles and potential of a new multidisciplinary tool. Trends Biotechnol 1996;14:261–72.

[2] Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 1992;89(22):10915–9.

[3] Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. Trends Genet 2000;16(6):276–7.

[4] Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 1970;48(3):443–53.

[5] Smith TF, Waterman MS. Identification of common molecular subsequences. J Mol Biol 1981;147(1):195–7.

[6] Higgins DG, Sharp PM. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 1988;73(1):237–44.

[7] Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 2000;302(1):205–17.

[8] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215(3):403–10.

[9] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25(17):3389–402.

[10] Crick F. Central dogma of molecular biology. Nature 1970;227(5258):561–3.

[11] Drake JW. Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci USA 1993;90:4171–5.

[12] Sarich VM, Wilson AC. Immunological time scale for hominid evolution. Science 1967;158(3805):1200–3.

[13] Woese C, Fox G. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 1977;74(11):5088–90.

[14] Ou CY, Ciesielski CA, Myers G, Bandea CI, Luo CC, Korber BT, et al. Molecular epidemiology of HIV transmission in a dental practice. Science 1992;256:1165–71.

[15] Shankarappa R, Gupta P, Learn Jr GH, Rodrigo AG, Rinaldo Jr CR, Gorry MC, et al. Evolution of human immunodeficiency virus type 1 envelope sequences in infected individuals with differing disease progression profiles. Virology 1998;241(2):251–9.

[16] The International Human Genome Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.

[17] Shankarappa R, Gupta P, Learn Jr GH, Rodrigo AG, Rinaldo Jr CR, Gorry MC, et al. Progress with methods for constructing evolutionary trees. Trends Ecol Evol 1992;7:73–9.

[18] Gramm J, Nickelsen A, Tantau T. Fixed-parameter algorithms in phylogenetics. Methods Mol Biol 2008;452:507–35.

[19] Collins DW, Jukes TH. rates of transition and transversion in coding sequences since the human-rodent divergence. Genomics 1994;20(3):386–96.

[20] Felsenstein J. Cases in which parsimony or compatibility methods will be positively misleading. Syst Zool 1978;27:401–10.

[21] Felsenstein J. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. Syst Zool 1973;22(3):240–9.

[22] Kidd KK, Sgaramella-Zonta LA. Phylogenetic analysis: concepts and methods. Am J Hum Genet 1971;23:235–52.

[23] Larget B, Simon LD. Markov Chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Mol Biol Evol 1999;16(6):750–9.

[24] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–9.

[25] Ling C, Hamada T, Bai J, Li X, Chesters D, Zheng W, et al. MrBayes tgMC^3: a tight GPU implementation of MrBayes. PLoS ONE 2013;8(4):e60667.

[26] Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol 1963;7:95–9.

[27] Protein data bank contents guide: atomic coordinate entry format description, version 3.3: 2011; July. Available at: <www.wwpdb.org/documentation/format33/v3.3.html>.

[28] <http://pymol.org/>.

[29] <http://jmol.sourceforge.net/>.

[30] Richardson JS. Early ribbon drawings of proteins. Nat Struct Biol 2000;7(8):624–5.

[31] Epstein CJ. Relation of protein evolution to tertiary structure. Nature 1964;203:1350–2.

[32] Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. EMBO J 1986;5(4):823–6.

[33] Sander C, Schneider R. Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins 1991;9(1):56–68.

[34] Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol 1990;183:63–98.

[35] Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 2003;31(13):3381–5.

[36] Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 2004;32(Web server issue):W526–31.

[37] Taylor WR. The classification of amino acid conservation. J Theor Biol 1986;119(2):205–18.

[38] Dodge C, Schneider R, Sander C. The HSSP database of protein structure-sequence alignments and family profiles. Nucleic Acids Res 1998;26(1):313–5.

[39] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;32(5):1792–7.

[40] Simons KT, Bonneau R, Ruczinski I, Baker D. Ab initio protein structure prediction of CASP III targets using ROSETTA. Proteins 1999(Suppl. 3):171–6.

[41] Fiser A, Do RK, Sali A. Modeling of loops in protein structures. Protein Sci 2000;9(9):1753–73.

[42] Adrian R, Ron EJ. Modeling side chains in peptides and proteins: application of the locally enhanced sampling and the simulated annealing methods to find minimum energy conformations. Chem Phys 1991;95:9277.

[43] Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol 1993;232(2):584–99.

[44] Bystroff C, Baker D. Prediction of local structure in proteins using a library of sequence-structure motifs. J Mol Biol 1998;281(3):565–77.

[45] Guo AM, Sun K, Su X, Wang H, Sun H. YY1TargetDB: an integral information resource for Yin Yang 1 target loci. Database 2013;2013:2013 [article ID bat007].

[46] Kodama Y, Shumway M, Leinonen ROn behalf of the International nucleotide sequence database collaboration. The sequence read archive: explosive growth of sequencing data. Nucleic Acids Res 2012;40(D1):D54–6.

[47] Nakamura Y, Cochrane G, Karsch-Mizrachi I The international nucleotide sequence database collaboration. Nucleic Acids Res 2011;41:D21–4.

[48] Zhu Y, Stephens RM, Meltzer PS, Davis SR. SRAdb: query and use public next generation sequencing data from within R. BMC Bioinformatics 2013;14:19.

[49] The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1092 human genomes. Nature 2012;491:56–65.

[50] Leinonen R, Akhtar R, Birney E, Bower L, Cerdeno-Tárraga A, Cheng Y, et al. The European nucleotide archive. Nucleic Acids Res 2011;39(suppl 1):D28–31.

[51] Cochrane G, Alako B, Amid C, Bower L, Cerdeño-Tárraga A, Cleland I, et al. Facing growth in the European nucleotide archive. Nucleic Acids Res 2013;41:D30–5.

[52] Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. Nucleic Acids Res 2013;10(D1):D36–42.

[53] National Center for Biotechnology Information. The NCBI Handbook [internet]. US: National Library of Medicine; 2002. Available from: www.ncbi.nlm.nih.gov/books/NBK21101.

[54] Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, et al. The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 2007;39:1181–6.

[55] Ludman EJ, Fullerton SM, Spangler L, Trinidad SB, Fujii MM, Jarvik GP, et al. Glad you asked: participants' opinions of re-consent for dbGaP data submission. J Empir Res Hum Res Ethics 2010;5:9–16.

[56] Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10.

[57] Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, et al. ArrayExpress—a public repository for microarray gene expression data at the EBI. Nucleic Acids Res 2003;31:68–71.

[58] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.

[59] Kogenaru S, Yan Q, Guo Y, Wang N. RNA-seq and microarray complement each other in transcriptome profiling. BMC Genomics 2012;13:629.

[60] Brazma A. Minimum information about a microarray experiment (MIAME)—successes, failures, challenges. Sci World J 2009;9:420–3.

[61] Functional Genomics Data Society. MINSEQE: minimum information about a high throughput nucleotide sequencing experiment a proposal for standards in functional genomic data reporting; 2012. Available from: www.fged.org/projects/minseqe/.

[62] Rustici G, Kolesnikov N, Brandizi M, Burdett T, Dylag M, Emam I, et al. ArrayExpress update—trends in database growth and links to data analysis tools. Nucleic Acids Res 2013;41:D987–90.

[63] Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol 2009;10:R130.

[64] Kapushesky M, Emam I, Holloway E, Kurnosov P, Zorin A, Malone J, et al. Gene expression atlas at the European bioinformatics institute. Nucleic Acids Res 2010;38:D690–8.

[65] Ohno S. So much "junk" DNA in our genome. Brookhaven Symp Biol 1972; 1972(23):366–70.

[66] de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet 2011;7:e1002384.

[67] Shapiro JA, von Sternberg R. Why repetitive DNA is essential to genome function. Biol Rev Camb Philos Soc 2005;80:227–50.

[68] Walker FO. Huntington's disease. Lancet 2007;369(9557):218–28.

[69] Jurka J, Kapitonov V, Pavlicek A, Klonowski P, Kohany O, et al. Repbase update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 2005;110:462–7.

[70] Coulombe Y, Lemieux M, Moreau J, Aubin J, Joksimovic M, Bérubé-Simard F, et al. Multiple promoters and alternative splicing: *Hoxa5* transcriptional complexity in the mouse embryo. PLoS ONE 2010;5:e10600.

[71] Périer RC, Praz V, Junier T, Bonnard C, Bucher P. The eukaryotic promoter database (EPD). Nucleic Acids Res 2000;14(1):302–3.

[72] Dreos R, Ambrosini G, Cavin Périer R, Bucher P. EPD and EPDnew, high-quality promoter resources in the next-generation sequencing era. Nucleic Acids Res 2013;41(Database issue):D157–64.

[73] Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 2006;34(Database issue):D108–10.

[74] Zhang HM, Chen H, Liu W, Liu H, Gong J, et al. AnimalTFDB: a comprehensive animal transcription factor database. Nucleic Acids Res 2012;40:D144–9.

[75] Chen K, Rajewsky N. The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet 2007;8:93–103.

[76] Yates LA, Norbury CJ, Gilbert RJ. The long and short of microRNA. Cell 2013;153(3):516–9.

[77] Griffiths-Jones S. MiRBase: the microRNA sequence database. Methods Mol Biol 2006;342:129–38.

[78] Dweep H, Sticht C, Pandey P, Gretz N. MiRWalk—database: prediction of possible miRNA binding sites by "walking" the genes of 3 genomes. J Biomed Inform 2011;2011(44):839–47.

[79] Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes Dev 2009;23:781–3.

[80] Fingerman Ian M, Zhang Xuan, Ratzat Walter, Husain Nora, Cohen Robert F, et al. NCBI epigenomics: what's new for 2013? Nucleic Acids Res 2013;41(Database issue):D221–5.

[81] The UniProt Consortium. Update on activities at the universal protein resource (UniProt) in 2013. Nucleic Acids Res 2013;41:D43–7.

[82] Vizcaino JA, Cote R, Csordas A, Dianes JA, et al. The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res 2013;41:D1063–9.

[83] Bernstein FC, Koetzle TF, Williams GJ, Meyer Jr EE, Brice MD, Rodgers JR, et al. The protein data bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535.

[84] Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0—the human metabolome database in 2013. Nucleic Acids Res 2013;41(Database issue):D801–7.

[85] Ogata H, Goto S, Fujibuchi W, Kanehisa M. Computation with the KEGG pathway database. Biosystems 1998;47(1–2):119–28.

[86] Altman T, Travers M, Kothari A, Caspi R, Karp PD. A systematic comparison of the MetaCyc and KEGG pathway databases. BMC Bioinformatics 2013;14:112.

[87] Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur O, Anwar N, et al. Pathway commons, a web resource for biological pathway data. Nucleic Acids Res 2011;39(Database issue):D685–90.

[88] Matthews L, Gopinath G, Gillespie M, Caudy M, Croft D, de Bono B, et al. Reactome knowledgebase of human biological pathways and processes. Nucleic Acids Res 2009;37(Database issue):D619–22.

[89] van Iersel MP, Kelder T, Pico AR, Hanspers K, Coort S, Conklin BR, et al. Presenting and exploring biological pathways with PathVisio. BMC Bioinformatics 2008;9:399.

[90] Kelder T, van Iersel MP, Hanspers K, Kutmon M, Conklin BR, Evelo C, et al. WikiPathways: building research communities on biological pathways. Nucleic Acids Res 2012;40(Database issue):D1301–7.

[91] Thorn CF, Klein TE, Altman RB, Thorn CF, Klein TE, Altman RB. PharmGKB: the pharmacogenomics knowledge base. Methods Mol Biol 2013;1015:311–20.

[92] The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. Nat Genet 2000;25(1):25–9.

[93] Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans institute of genetic medicine, Johns Hopkins University, 2013. http://omim.org/.

[94] Scott F. The NCBI taxonomy database. Nucleic Acids Res 2012;40(Database issue):D136–43.

[95] Codd EF. A relational model of data for large shared data banks. Commun ACM 1970;13(6):377–87.

[96] Cock PJ, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. Nucleic Acids Res 2010;38(6):1767–71.

This page is intentionally left blank

Chapter 14

PharmGKB: The Pharmacogenomics Knowledgebase

Julia M. Barbarino,[a] **Michelle Whirl-Carrillo,**[a] **Teri E. Klein,**[a] **and Russ B. Altman**[b]

[a]*Department of Genetics, Stanford University, USA,* [b]*Department of Genetics, Department of Bioengineering, Stanford University, USA*

Chapter Outline

14.1 Overview 289
14.2 The PharmGKB Website 290
14.3 Variant Annotations 291
14.4 Clinical Annotations 294
14.5 VIPs and Pathways 297
14.6 Clinical Implementation: Consortia, CPIC, and Dosing Guidelines 301
14.7 Projects in Development 303
14.8 Conclusion 304
Glossary 305
Acronyms and Abbreviations 305
Acknowledgments 306
References 306

14.1 OVERVIEW

The field of pharmacogenomics has grown rapidly over the past several decades, expanding alongside the larger revolution in personalized medicine. This expansion has both led to, and been driven by, an increase in research identifying gene–drug relationships. These associations and relationships are published in diverse articles in a number of journals, and they represent significant additions to our understanding of both pharmacogenetics and pharmacogenomics (PGx).* PharmGKB (www.pharmgkb.org) is funded by the National Institutes of Health (NIH) and the National Institute of General Medical Sciences (NIGMS), and includes a publicly available online knowledgebase that

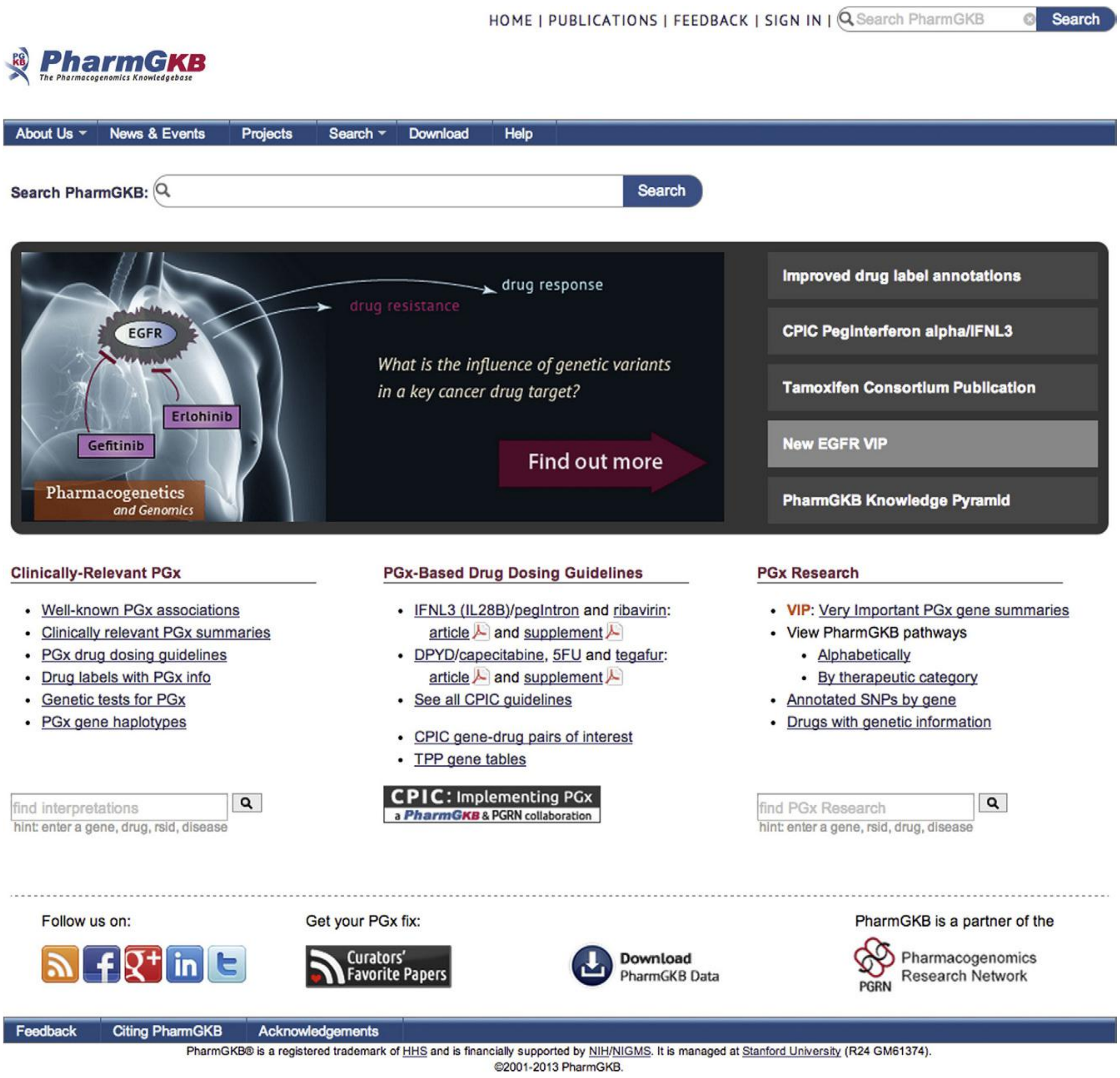
1. Annotates, aggregates, and integrates reported study results and stores this information in a systematic and useful manner.
2. Contributes to clinical pharmacogenomic implementation projects.
3. Enables consortia, both data-centric and knowledge-centric, to address pharmacogenomic questions and develop drug-dosing guidelines.

The foundation of PharmGKB consists of its vast collection of PGx study results, and it is a valuable resource for clinicians and researchers across broad areas of science and medicine. This wealth of data also places PharmGKB in the unique position of being able to recognize clinically relevant results. Creating high-level annotations that identify variant–drug associations with a preponderance of supporting evidence provides the underlying knowledge needed for clinical pharmacogenomic implementation. PharmGKB's mission goes beyond the acquisition and dissemination of knowledge to include co-leading the Clinical Pharmacogenetics Implementation Consortium (CPIC), which provides drug-dosing guidelines based on an individual's genotype. This chapter describes the variety of knowledge and knowledge-based activities available in PharmGKB and how the knowledgebase can be easily used to find out more about PGx associations of interest.

---

* Pharmacogenetics is the study of genetic influence on drug response, usually involving one or two genes. Pharmacogenomics is the study of how genomic variation influences drug response across the genome. Although technically distinct, we use the term *pharmacogenomics* (PGx) throughout this chapter to represent both.

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00014-1
© 2014 Elsevier Inc. All rights reserved.


HOME | PUBLICATIONS | FEEDBACK | SIGN IN |
Search PharmGKB
Search
PharmGKB
The Pharmacogenomics Knowledgebase
About Us
News & Events
Projects
Search
Download
Help
Search PharmGKB:
Search
drug response
drug resistance
EGFR
Erlohinib
Gefitinib
Pharmacogenetics and Genomics
Find out more
Improved drug label annotations
CPIC Peginterferon alpha/IFNL3
Tamoxifen Consortium Publication
New EGFR VIP
PharmGKB Knowledge Pyramid
Clinically-Relevant PGx
Well-known PGx associations
Clinically relevant PGx summaries
PGx drug dosing guidelines
Drug labels with PGx info
Genetic tests for PGx
PGx gene haplotypes
PGx-Based Drug Dosing Guidelines
IFNL3 (IL28B)/pegIntron and ribavirin:
article and supplement
DPYD/capecitabine, 5FU and tegafur:
article and supplement
See all CPIC guidelines
CPIC gene-drug pairs of interest
TPP gene tables
PGx Research
VIP: Very Important PGx gene summaries
View PharmGKB pathways
Alphabetically
By therapeutic category
Annotated SNPs by gene
Drugs with genetic information
find interpretations
hint: enter a gene, drug, rsid, disease
CPIC: Implementing PGx
a PharmGKB & PGRN collaboration
find PGx Research
hint: enter a gene, rsid, drug, disease
Follow us on:
Get your PGx fix:
Curators' Favorite Papers
Download PharmGKB Data
PharmGKB is a partner of the
PGRN Pharmacogenomics Research Network
Feedback
Citing PharmGKB
Acknowledgements

What is the influence of genetic variants in a key cancer drug target?

PharmGKB® is a registered trademark of HHS and is financially supported by NIH/NIGMS. It is managed at Stanford University (R24 GM61374).
©2001-2013 PharmGKB.

FIGURE 14.1 **PharmGKB homepage.** From the PharmGKB homepage, all knowledge on PharmGKB can be easily accessed, either by searching for a variant, gene, drug, disease, or phenotype of interest, or through direct links provided at the bottom of the page. Under the "Clinically-Relevant PGx summaries" heading on the left, links are provided to high-level clinical annotations, drug-dosing guidelines, and drug labels. Under "PGx Research" on the right, links are provided to VIP summaries and pathways. A rotating box in the center provides updates on the latest information added to the website.

14.2 THE PHARMGKB WEBSITE

The PharmGKB website contains several principal types of knowledge, which are discussed in the following sections. These include variant annotations, clinical annotations, very important pharmacogene (VIP) summaries, pathways, dosing guidelines, and drug labels. The PharmGKB homepage can be seen in Figure 14.1. From this page, users can search for variants, genes, drugs, and diseases or phenotypes of interest, or access the different types of PharmGKB knowledge through direct links at the bottom of the page.

A visual summary of the various types of information and how they are integrated in PharmGKB is provided in the PharmGKB Knowledge Pyramid (Figure 14.2). Primary PGx literature represents the basis of all knowledge created on PharmGKB. Through both manual curation and natural-language processing (NLP) (see Section 14.7 for more on the latter), PharmGKB extracts information from this literature and provides it on the website in the form of variant annotations, where each variant annotation describes an association between a single genetic variant and a drug phenotype as found in a publication. Using the

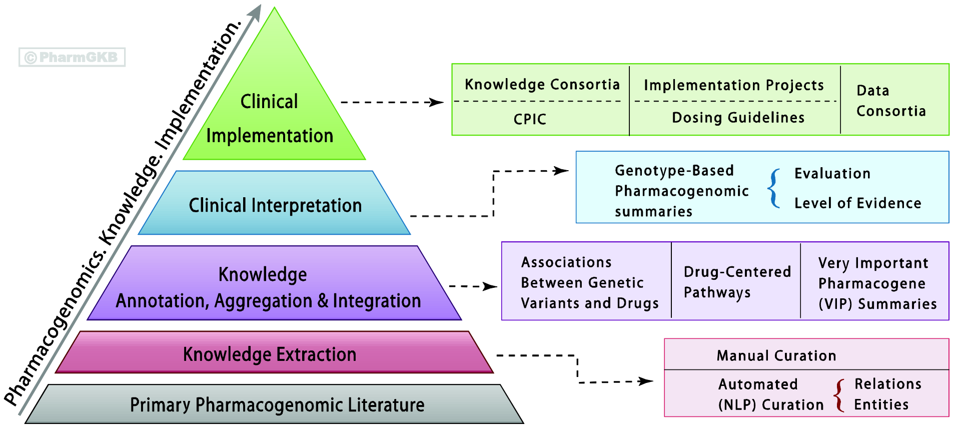

©PharmGKB
Pharmacogenomics. Knowledge. Implementation.
Clinical Implementation
Clinical Interpretation
Knowledge Annotation, Aggregation & Integration
Knowledge Extraction
Primary Pharmacogenomic Literature
Knowledge Consortia
CPIC
Implementation Projects
Dosing Guidelines
Data Consortia
Genotype-Based Pharmacogenomic summaries
Evaluation
Level of Evidence
Associations Between Genetic Variants and Drugs
Drug-Centered Pathways
Very Important Pharmacogene (VIP) Summaries
Manual Curation
Automated (NLP) Curation
Relations
Entities

FIGURE 14.2 **The PharmGKB Knowledge Pyramid.** The PharmGKB Knowledge Pyramid is a visual representation of the various types of information available on PharmGKB. Primary pharmacogenomics literature forms the base, from which PharmGKB curators extract relevant PGx information ("Knowledge Extraction") and create variant annotations, VIP summaries, and pathways ("Knowledge Annotation, Aggregation & Integration"). Clinical annotations build on variant annotations ("Clinical Interpretation"). Clinical implementation, shown at the top of the pyramid, represents an important end goal for PharmGKB.

primary literature information, PharmGKB curators can also create original materials such as VIP summaries and drug-centered pathways. Clinical annotations build further on variant annotations and are clinically inclined and evidence-rated summaries of all variant annotations that discuss a particular variant and drug phenotype. PharmGKB's involvement in clinical implementation projects and data-centric consortia is represented at the top level of the pyramid, as implementation is supported by all other knowledge in PharmGKB. Through the creation of its various materials, PharmGKB is able to push PGx information upward from original publications toward potential clinical implementation.

Throughout this chapter, we primarily use the *CYP3A5* single-nucleotide polymorphism (SNP) rs776746 (also known as *CYP3A5**3) as an example variant, as well as the immunosuppressant tacrolimus as an example drug, to illustrate the types of knowledge a user can ascertain from the PharmGKB website. The rs776746 C allele results in a nonfunctional CYP3A5 protein [1] and since CYP3A5 is responsible for metabolizing tacrolimus, [2] this variant and drug represent a PGx association of interest.

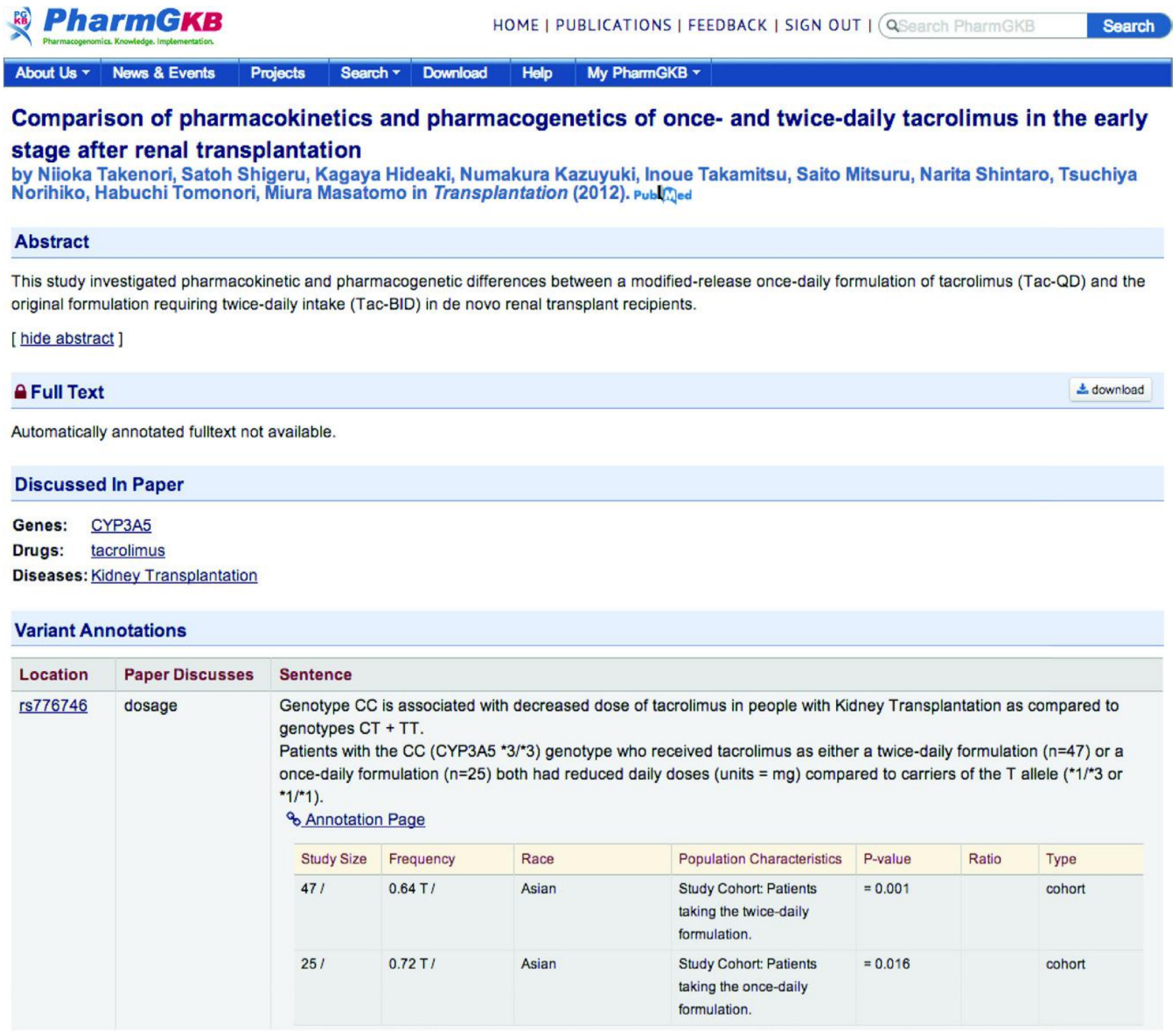
14.3 VARIANT ANNOTATIONS

Variant annotations are core components of PharmGKB; they are summaries describing how a particular genetic variant is associated with a drug response, such as efficacy or toxicity, as reported in a single publication. To create these summaries, PharmGKB curators manually review past and current literature and add any studies containing relevant pharmacogenomic information to the knowledgebase. While PharmGKB aims to capture as many studies as possible, complete coverage is not feasible given the volume of literature and resources needed for manual curation. However, ongoing work using NLP will greatly aid these efforts in the future (see Section 14.7).

An example variant annotation is shown in Figure 14.3. Information about the publication, including an abstract, appears at the top of the page and annotations appear at the bottom. In this particular example, the publication of interest is a 2012 study by Niioka et al. [3]. Here, the authors found that the daily dose of tacrolimus was lower in patients with the rs776746 CC genotype as compared to those with the CT or TT genotype. Since this association is a relevant PGx finding, a variant annotation was created for it.

Variant annotations can be created for positive or negative results and consist of a summary describing the study results as well as other important information such as study size, allele frequencies, p-values, and other statistical results. Annotations are written for variants other than SNPs, including copy number variants (CNVs) and insertions and deletions (indels). Additionally, they can be submitted on haplotypes, and full haplotype spreadsheets for certain genes are available for download on the PharmGKB site. The term "variant" is used throughout this chapter to indicate these different types of variants as well as haplotypes. PharmGKB includes annotations on many primary studies, including genome-wide association studies (GWAS), clinical trials, and functional *in vitro* studies. Additional information about variant annotations can be found in the review by McDonagh et al. [4].

For brevity, only one annotation is shown in Figure 14.3. However, Niioka et al. found other pharmacokinetic results in their study that did not directly pertain to dose. These were added as additional, separate annotations and appear below the example annotation on the original webpage. A research article annotated in PharmGKB can have as many



Comparison of pharmacokinetics and pharmacogenetics of once- and twice-daily tacrolimus in the early stage after renal transplantation

by Niioka Takenori, Satoh Shigeru, Kagaya Hideaki, Numakura Kazuyuki, Inoue Takamitsu, Saito Mitsuru, Narita Shintaro, Tsuchiya Norihiko, Habuchi Tomonori, Miura Masatomo in *Transplantation* (2012).

Abstract

This study investigated pharmacokinetic and pharmacogenetic differences between a modified-release once-daily formulation of tacrolimus (Tac-QD) and the original formulation requiring twice-daily intake (Tac-BID) in de novo renal transplant recipients.

[hide abstract]

Full Text

Automatically annotated fulltext not available.

Discussed In Paper

Genes: CYP3A5
Drugs: tacrolimus
Diseases: Kidney Transplantation

Variant Annotations

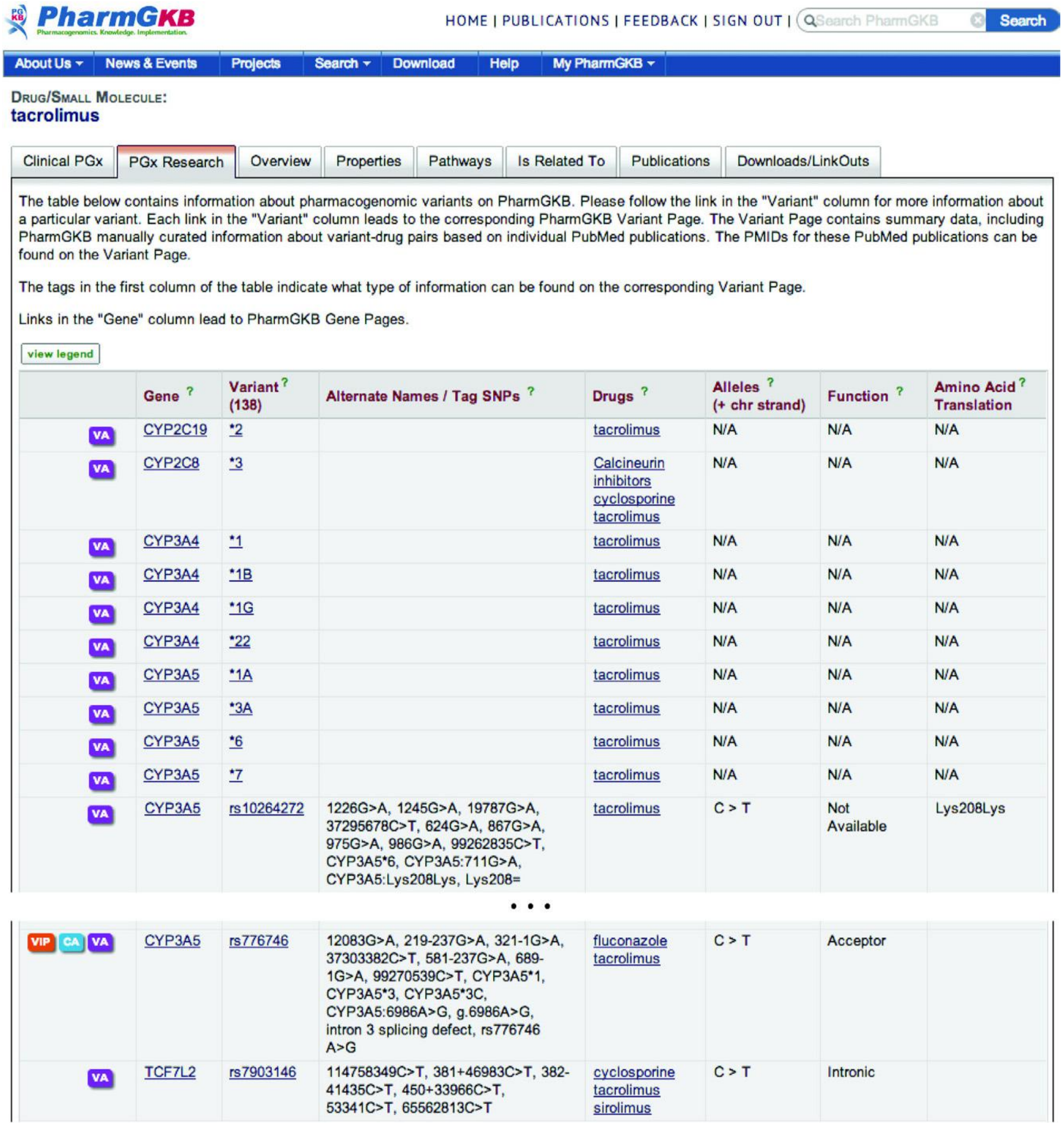
Location	Paper Discusses	Sentence
rs776746	dosage	Genotype CC is associated with decreased dose of tacrolimus in people with Kidney Transplantation as compared to genotypes CT + TT. Patients with the CC (CYP3A5 *3/*3) genotype who received tacrolimus as either a twice-daily formulation (n=47) or a once-daily formulation (n=25) both had reduced daily doses (units = mg) compared to carriers of the T allele (*1/*3 or *1/*1). Annotation Page

Study Size	Frequency	Race	Population Characteristics	P-value	Ratio	Type
47 /	0.64 T /	Asian	Study Cohort: Patients taking the twice-daily formulation.	= 0.001		cohort
25 /	0.72 T /	Asian	Study Cohort: Patients taking the once-daily formulation.	= 0.016		cohort

FIGURE 14.3 **PharmGKB variant annotation.** Annotations are created on individual publications and consist of standardized summaries of any pharmacogenomic results found in the published study results. Multiple annotations can be created for a single study if multiple associations are found. Information about the publication, including an abstract, is shown at the top of the page; annotations appear at the bottom. The full annotation can be viewed at www.pharmgkb.org/pmid/23073468.

annotations as the different associations found. The enormous number of annotated publications on the PharmGKB website, combined with the varied number of individual annotations on each publication, leads to a large amount of data, which PharmGKB keeps organized using several methods. Every variant annotation contains one summary sentence written in a highly standardized format; in Figure 14.3 that sentence is the first one under the "Sentence" column in the Variant Annotations section. Every annotation added to PharmGKB is entered using a curator annotation tool developed by the PharmGKB staff, restricting text to a predefined format. This allows results to be easily compared and contrasted between or within studies. However, space for free text is provided below the standardized sentence, in case further elaboration on results is necessary (as seen in Figure 14.3, where the sentence in the second paragraph is free text).

PharmGKB also keeps its information organized by storing and structuring data around four specific entities: variants, genes, drugs, and diseases or phenotypes. Every variant annotation contains a variant, and most will include drug information. Annotations also include a disease or phenotype if appropriate; examples include Stevens-Johnson syndrome, which is often seen in annotations describing



DRUG/SMALL MOLECULE:
tacrolimus

The table below contains information about pharmacogenomic variants on PharmGKB. Please follow the link in the "Variant" column for more information about a particular variant. Each link in the "Variant" column leads to the corresponding PharmGKB Variant Page. The Variant Page contains summary data, including PharmGKB manually curated information about variant-drug pairs based on individual PubMed publications. The PMIDs for these PubMed publications can be found on the Variant Page.

The tags in the first column of the table indicate what type of information can be found on the corresponding Variant Page.

Links in the "Gene" column lead to PharmGKB Gene Pages.

	Gene ?	Variant ? (138)	Alternate Names / Tag SNPs ?	Drugs ?	Alleles ? (+ chr strand)	Function ?	Amino Acid ? Translation
VA	CYP2C19	*2		tacrolimus	N/A	N/A	N/A
VA	CYP2C8	*3		Calcineurin inhibitors cyclosporine tacrolimus	N/A	N/A	N/A
VA	CYP3A4	*1		tacrolimus	N/A	N/A	N/A
VA	CYP3A4	*1B		tacrolimus	N/A	N/A	N/A
VA	CYP3A4	*1G		tacrolimus	N/A	N/A	N/A
VA	CYP3A4	*22		tacrolimus	N/A	N/A	N/A
VA	CYP3A5	*1A		tacrolimus	N/A	N/A	N/A
VA	CYP3A5	*3A		tacrolimus	N/A	N/A	N/A
VA	CYP3A5	*6		tacrolimus	N/A	N/A	N/A
VA	CYP3A5	*7		tacrolimus	N/A	N/A	N/A
VA	CYP3A5	rs10264272	1226G>A, 1245G>A, 19787G>A, 37295678C>T, 624G>A, 867G>A, 975G>A, 986G>A, 99262835C>T, CYP3A5*6, CYP3A5:711G>A, CYP3A5:Lys208Lys, Lys208=	tacrolimus	C > T	Not Available	Lys208Lys
• • •							
VIP CA VA	CYP3A5	rs776746	12083G>A, 219-237G>A, 321-1G>A, 37303382C>T, 581-237G>A, 689-1G>A, 99270539C>T, CYP3A5*1, CYP3A5*3, CYP3A5*3C, CYP3A5:6986A>G, g.6986A>G, intron 3 splicing defect, rs776746 A>G	fluconazole tacrolimus	C > T	Acceptor	
VA	TCF7L2	rs7903146	114758349C>T, 381+46983C>T, 382-41435C>T, 450+33966C>T, 53341C>T, 65562813C>T	cyclosporine tacrolimus sirolimus	C > T	Intronic	

FIGURE 14.4 **PharmGKB drug page for tacrolimus.** Different tabs at the top under "tacrolimus" provide different information about this drug. The "PGx Research" tab provides a table of variants and corresponding genes that have been associated with that drug in variant annotations. Colorful icons in the far left column indicate whether other available PharmGKB resources discuss that particular variant and a drug, such as VIPs (orange "VIP") or clinical annotations (blue "CA"); clicking "view legend" explains the meaning of each icon. The "Clinical PGx" tab provides information about the drug from other PharmGKB sources, such as clinical annotations or dosing guidelines. Other tabs provide additional information about the drug ("Overview"; "Properties") and lists of publications with annotations that include it ("Publications").

an adverse reaction to allopurinol in *HLA-B**58:01 carriers, and neutropenia, which is typically used in annotations describing side effects of irinotecan in *UGT1A1**28 carriers. Standardized vocabularies are used for these entities, ensuring regularization of terms across all annotations. PharmGKB sources several outside databases for these vocabularies. Variant annotations are typically created for a specific reference SNP identification (rsID) as defined by the National Center for Biotechnology Information (NCBI) SNP database (dbSNP) [5]. For haplotypes, PharmGKB curates information from the literature or from nomenclature resources such as the Human Cytochrome P450 (CYP) Allele Nomenclature Database [6]; the UDP-Glucuronosyltransferase Alleles Nomenclature page [7]; and the HLA Nomenclature resource [8]. Gene symbols come from the HUGO Gene Nomenclature Committee (HGNC) [9]; drug, compound, or substance information comes from DrugBank [10], WHO's Anatomical Therapeutic Chemical (ATC) classification system, [11] or PubChem [12]. Disease and phenotype ontologies are taken from the National Library of Medicine's Medical Subject Headings (MeSH) browser [13].

PharmGKB curators can add terms to each of these vocabularies or ontologies if necessary. rsIDs, genes, drugs and diseases, or phenotypes present in these pre-existing vocabularies have their own pages on PharmGKB. When a variant annotation is created that includes entities from these sources, the relevant page is populated with information from that annotation. Additionally, each time entities in these vocabularies are discussed together in a variant annotation, they are stored as a relationship in the database. This creates a network of connections throughout PharmGKB that can be used for further research and discovery. PharmGKB is currently working on a visualization prototype for this data (see Section 14.7), but any registered user can download a file of these relationships from PharmGKB and use them in further research.

Organizing information around the four entities makes finding information on PharmGKB straightforward. Users can search for a gene, drug, disease, or variant of interest and be taken to its specific page, where all PharmGKB knowledge on the entity is accumulated. This includes variant annotation data as well as information from clinical annotations, VIP summaries, pathways, dosing guidelines, and drug labels. Information from variant annotations is amassed under the "PGx Research" tab on a variant, gene, drug, or disease page. As an example, the drug page for tacrolimus is shown in Figure 14.4. A table lists the variants and corresponding genes that have been associated with that drug in variant annotations, and rs776746 and tacrolimus can be seen near the bottom of the figure. Colorful icons in the far left column indicate whether a variant and drug appear together in other PharmGKB resources, such as a clinical annotation (indicated by the blue "CA" icon) or a VIP (the orange "VIP" icon). Icons also exist for pathways, dosing guidelines, and drug labels, and more information about them can be found by clicking the "view legend" button on any drug, gene or disease page.

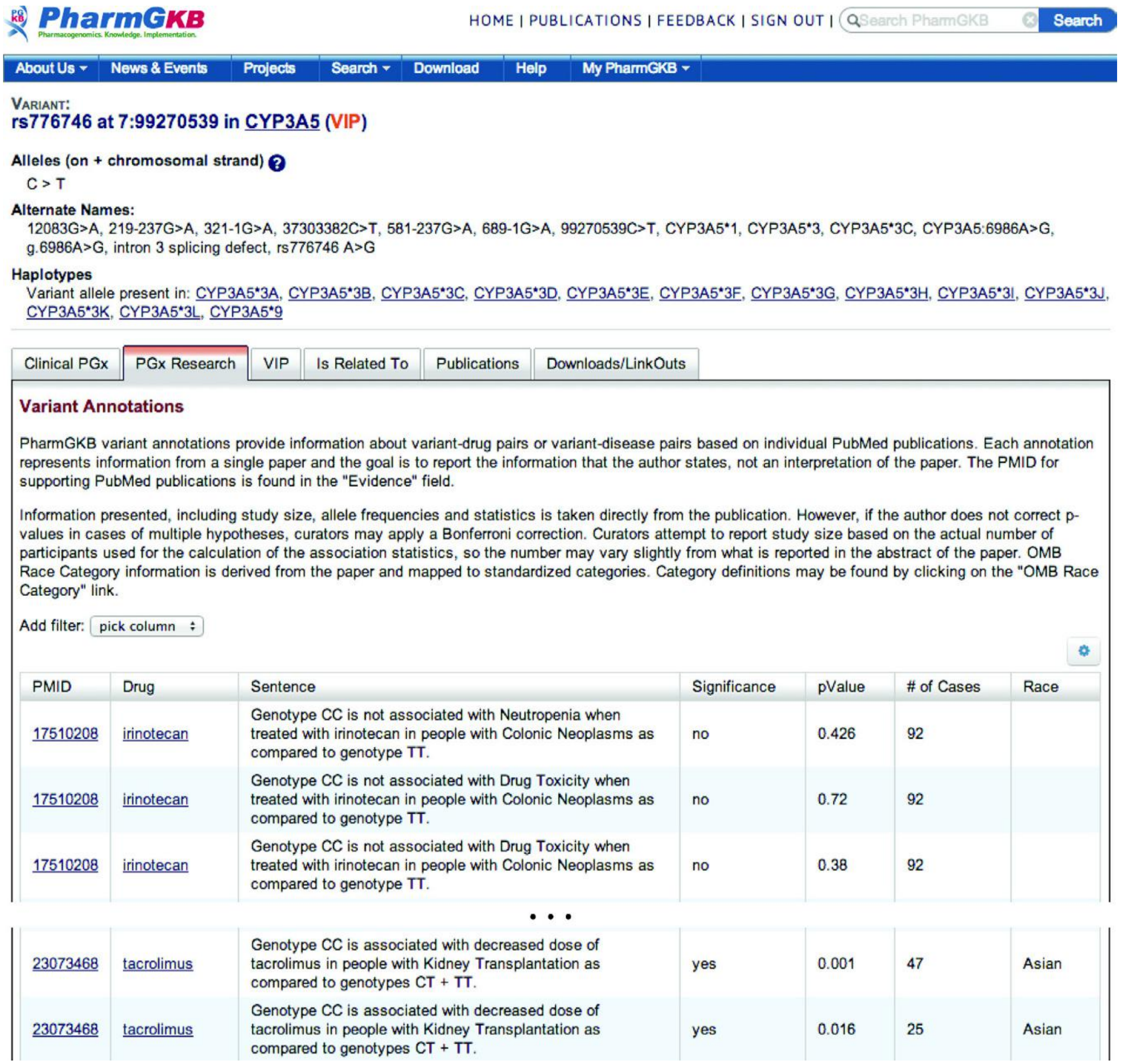
Gene pages are similar to drug pages, except that only variants from that gene are listed and all associated drugs are shown. Disease or phenotype pages can be extensive since they are not filtered by a drug or gene, and they list every annotation that mentions the disease or phenotype. Variant pages are slightly different, providing a table of all publications that have annotations discussing that variant, organized by PubMed ID (PMID). An image of the variant page for rs776746 can be seen in Figure 14.5; the annotations on the study by Niioka et al. (PMID 23073468) can be seen at the bottom of the figure. The table provides information from each annotation, including associated drug, standardized sentence, statistical significance, p-value, study size, and race. It is also filterable and sortable. Clicking on the PMID link opens the publication page to view the original annotation.

PharmGKB variant annotations can be used individually, such as by students interested in a particular PGx study, or in bulk, such as by researchers who wish to conduct further pharmacogenomic analyses. In addition to their consistency, quality, and volume, variant annotations are also valuable because they are easily aggregated. Annotations that discuss the same variant and drug response can be readily accumulated, allowing for assessment of clinical relevance. This is important from a clinical implementation perspective, and PharmGKB takes advantage of this ability through the creation of clinical annotations.

14.4 CLINICAL ANNOTATIONS

Clinical annotations are summaries of PharmGKB's literature evidence for a particular variant–drug association. Recall that every page for a specific variant includes a table of all the publications that include the variant in an annotation, as well as the associated drug and other relevant information. A user can look at this page and filter the table to see whether there are consistent associations between a variant and a drug or class of drugs. However, PharmGKB condenses this information in a concise, evidence-rated summary known as a clinical annotation. Clinical annotations combine all of the studies that show a relationship between a variant and a drug phenotype (or closely related phenotypes) and bring them together on one page. Each clinical annotation provides a short summary of the relationship for each genotype (or haplotype) compared to other genotypes, as well as a list of every study used as evidence for the association; this allows users to view the exact results supporting a clinical annotation.

More information on clinical annotations, as well as on why PharmGKB reports relationships relative to other genotypes, can be found in the paper by Whirl-Carrillo et al. [14]. Multiple studies have found an association between rs776746 and tacrolimus dose and metabolism; the clinical annotation for this association can be seen in Figure 14.6. The genotype-based summary is at the top of the page, and a selection of the evidence list is at the bottom; the study by Niioka et al. is evidence number 10. Metabolism and dose are closely linked, but studies that discuss disparate phenotypes receive separate clinical annotations. For instance, rs776746 and tacrolimus have additional clinical annotations for associations with nephrotoxicity and transplant rejection.



VARIANT:
rs776746 at 7:99270539 in CYP3A5 (VIP)

Alleles (on + chromosomal strand)
C > T

Alternate Names:
12083G>A, 219-237G>A, 321-1G>A, 37303382C>T, 581-237G>A, 689-1G>A, 99270539C>T, CYP3A5*1, CYP3A5*3, CYP3A5*3C, CYP3A5:6986A>G, g.6986A>G, intron 3 splicing defect, rs776746 A>G

Haplotypes
Variant allele present in: CYP3A5*3A, CYP3A5*3B, CYP3A5*3C, CYP3A5*3D, CYP3A5*3E, CYP3A5*3F, CYP3A5*3G, CYP3A5*3H, CYP3A5*3I, CYP3A5*3J, CYP3A5*3K, CYP3A5*3L, CYP3A5*9

Variant Annotations

PharmGKB variant annotations provide information about variant-drug pairs or variant-disease pairs based on individual PubMed publications. Each annotation represents information from a single paper and the goal is to report the information that the author states, not an interpretation of the paper. The PMID for supporting PubMed publications is found in the "Evidence" field.

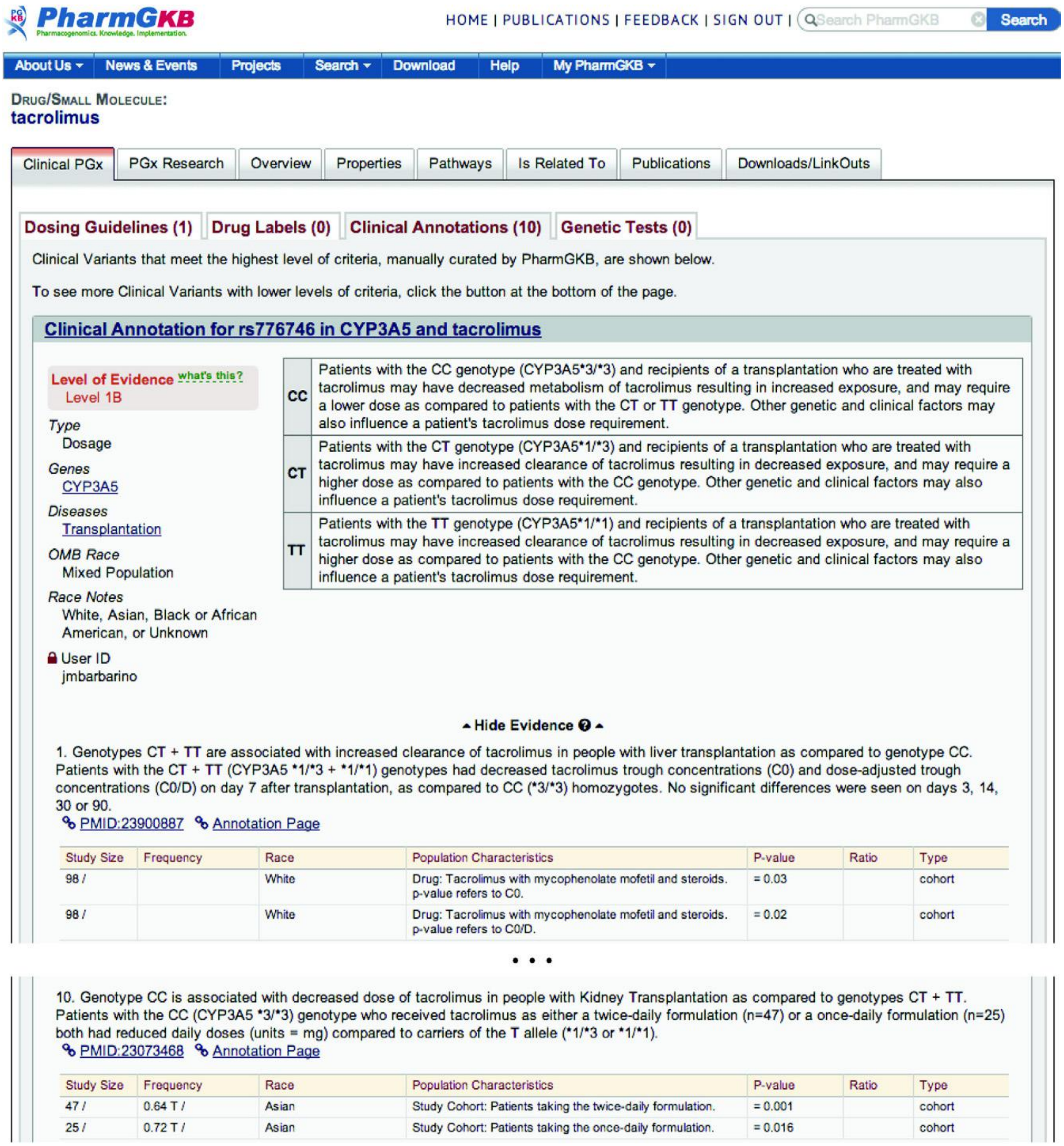
Information presented, including study size, allele frequencies and statistics is taken directly from the publication. However, if the author does not correct p-values in cases of multiple hypotheses, curators may apply a Bonferroni correction. Curators attempt to report study size based on the actual number of participants used for the calculation of the association statistics, so the number may vary slightly from what is reported in the abstract of the paper. OMB Race Category information is derived from the paper and mapped to standardized categories. Category definitions may be found by clicking on the "OMB Race Category" link.

PMID	Drug	Sentence	Significance	pValue	# of Cases	Race
17510208	irinotecan	Genotype CC is not associated with Neutropenia when treated with irinotecan in people with Colonic Neoplasms as compared to genotype TT.	no	0.426	92	
17510208	irinotecan	Genotype CC is not associated with Drug Toxicity when treated with irinotecan in people with Colonic Neoplasms as compared to genotype TT.	no	0.72	92	
17510208	irinotecan	Genotype CC is not associated with Drug Toxicity when treated with irinotecan in people with Colonic Neoplasms as compared to genotype TT.	no	0.38	92	
• • •						
23073468	tacrolimus	Genotype CC is associated with decreased dose of tacrolimus in people with Kidney Transplantation as compared to genotypes CT + TT.	yes	0.001	47	Asian
23073468	tacrolimus	Genotype CC is associated with decreased dose of tacrolimus in people with Kidney Transplantation as compared to genotypes CT + TT.	yes	0.016	25	Asian

FIGURE 14.5 **Variant page for rs776746.** Information about the variant, such as alleles and alternate names, is provided at the top of a variant annotation page. Under the "PGx Research" tab, a table is provided that lists each publication (by PMID) that has an annotation using that variant. Additional information from the study is also provided, such as p-value and study size. Clicking on the PMID opens the original annotation. Additional tabs provide information similar to that discussed in Figure 14.4. Here only five annotations that include rs776746 are shown, although many more exist for this variant. See www.pharmgkb.org/rsid/rs776746 for a full list.

After writing the summaries and viewing the supporting evidence, curators assign every clinical annotation a level of evidence, giving a measure of the confidence in the association. Levels of evidence are based on several criteria, including study sizes, p-values, odds ratios, and replication of the association. Negative results also play a critical role in this determination. An overview of the meaning behind the various levels of evidence can be seen in Figure 14.7 and is discussed in the paper by Whirl-Carrillo et al. [14]. This information can also be accessed by clicking the green "what's this" button next to the evidence level (see Figure 14.6).

Many variant annotations on PharmGKB are singular, meaning that there is no replication of an association between that genetic variant and a drug response. These studies are assigned an evidence level of 3. Clinical



DRUG/SMALL MOLECULE:
tacrolimus

Clinical Variants that meet the highest level of criteria, manually curated by PharmGKB, are shown below.

To see more Clinical Variants with lower levels of criteria, click the button at the bottom of the page.

Clinical Annotation for rs776746 in CYP3A5 and tacrolimus

Level of Evidence
Level 1B

Type
Dosage

Genes
CYP3A5

Diseases
Transplantation

OMB Race
Mixed Population

Race Notes
White, Asian, Black or African American, or Unknown

User ID
jmbarbarino

CC	Patients with the CC genotype (CYP3A5*3/*3) and recipients of a transplantation who are treated with tacrolimus may have decreased metabolism of tacrolimus resulting in increased exposure, and may require a lower dose as compared to patients with the CT or TT genotype. Other genetic and clinical factors may also influence a patient's tacrolimus dose requirement.
CT	Patients with the CT genotype (CYP3A5*1/*3) and recipients of a transplantation who are treated with tacrolimus may have increased clearance of tacrolimus resulting in decreased exposure, and may require a higher dose as compared to patients with the CC genotype. Other genetic and clinical factors may also influence a patient's tacrolimus dose requirement.
TT	Patients with the TT genotype (CYP3A5*1/*1) and recipients of a transplantation who are treated with tacrolimus may have increased clearance of tacrolimus resulting in decreased exposure, and may require a higher dose as compared to patients with the CC genotype. Other genetic and clinical factors may also influence a patient's tacrolimus dose requirement.

1. Genotypes CT + TT are associated with increased clearance of tacrolimus in people with liver transplantation as compared to genotype CC. Patients with the CT + TT (CYP3A5 *1/*3 + *1/*1) genotypes had decreased tacrolimus trough concentrations (C0) and dose-adjusted trough concentrations (C0/D) on day 7 after transplantation, as compared to CC (*3/*3) homozygotes. No significant differences were seen on days 3, 14, 30 or 90.
PMID:23900887 Annotation Page

Study Size	Frequency	Race	Population Characteristics	P-value	Ratio	Type
98 /		White	Drug: Tacrolimus with mycophenolate mofetil and steroids. p-value refers to C0.	= 0.03		cohort
98 /		White	Drug: Tacrolimus with mycophenolate mofetil and steroids. p-value refers to C0/D.	= 0.02		cohort

• • •

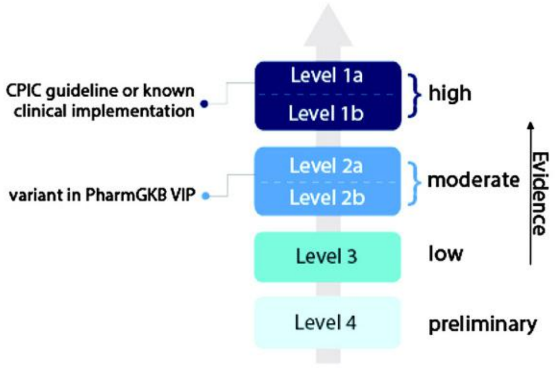
10. Genotype CC is associated with decreased dose of tacrolimus in people with Kidney Transplantation as compared to genotypes CT + TT. Patients with the CC (CYP3A5 *3/*3) genotype who received tacrolimus as either a twice-daily formulation (n=47) or a once-daily formulation (n=25) both had reduced daily doses (units = mg) compared to carriers of the T allele (*1/*3 or *1/*1).
PMID:23073468 Annotation Page

Study Size	Frequency	Race	Population Characteristics	P-value	Ratio	Type
47 /	0.64 T /	Asian	Study Cohort: Patients taking the twice-daily formulation.	= 0.001		cohort
25 /	0.72 T /	Asian	Study Cohort: Patients taking the once-daily formulation.	= 0.016		cohort

FIGURE 14.6 **Clinical annotation for rs776746 in *CYP3A5* and tacrolimus.** Clinical annotations can be found under the "Clinical PGx" and "Clinical Annotation" tabs on the variant, gene, or drug page. They consist of short genotype-based summaries of the literature evidence behind a variant–drug association. All clinical annotations are given a level of evidence that indicates the strength of the association. Supporting evidence behind the literature summary is given in a drop-down list below the text.

annotations are not created if associations between a variant and drug phenotype are exclusively negative. However, as more literature is published, additional studies may accumulate that discuss that particular association, with either positive or negative results.

PharmGKB curators periodically reassess the level of evidence of clinical annotations, adding recently created variant annotations as support. Variant–drug interactions can therefore move up or down through the levels, giving an indication of whether an association may be clinically



Level 1A

Annotation for a variant-drug combination in a CPIC or medical society-endorsed PGx guideline, or implemented at a PGRN site or in another major health system.

Level 1B

Annotation for a variant-drug combination where the preponderance of evidence shows an association. The association must be replicated in more than one cohort with significant p-values, and preferably will have a strong effect size.

Level 2A

Annotation for a variant-drug combination that qualifies for level 2B where the variant is within a VIP (Very Important Pharmacogene) as defined by PharmGKB. The variants in level 2A are in known pharmacogenes, so functional significance is more likely.

Level 2B

Annotation for a variant-drug combination with moderate evidence of an association. The association must be replicated but there may be some studies that do not show statistical significance, and/or the effect size may be small.

Level 3

Annotation for a variant-drug combination based on a single significant (not yet replicated) or annotation for a variant-drug combination evaluated in multiple studies but lacking clear evidence of an association.

Level 4

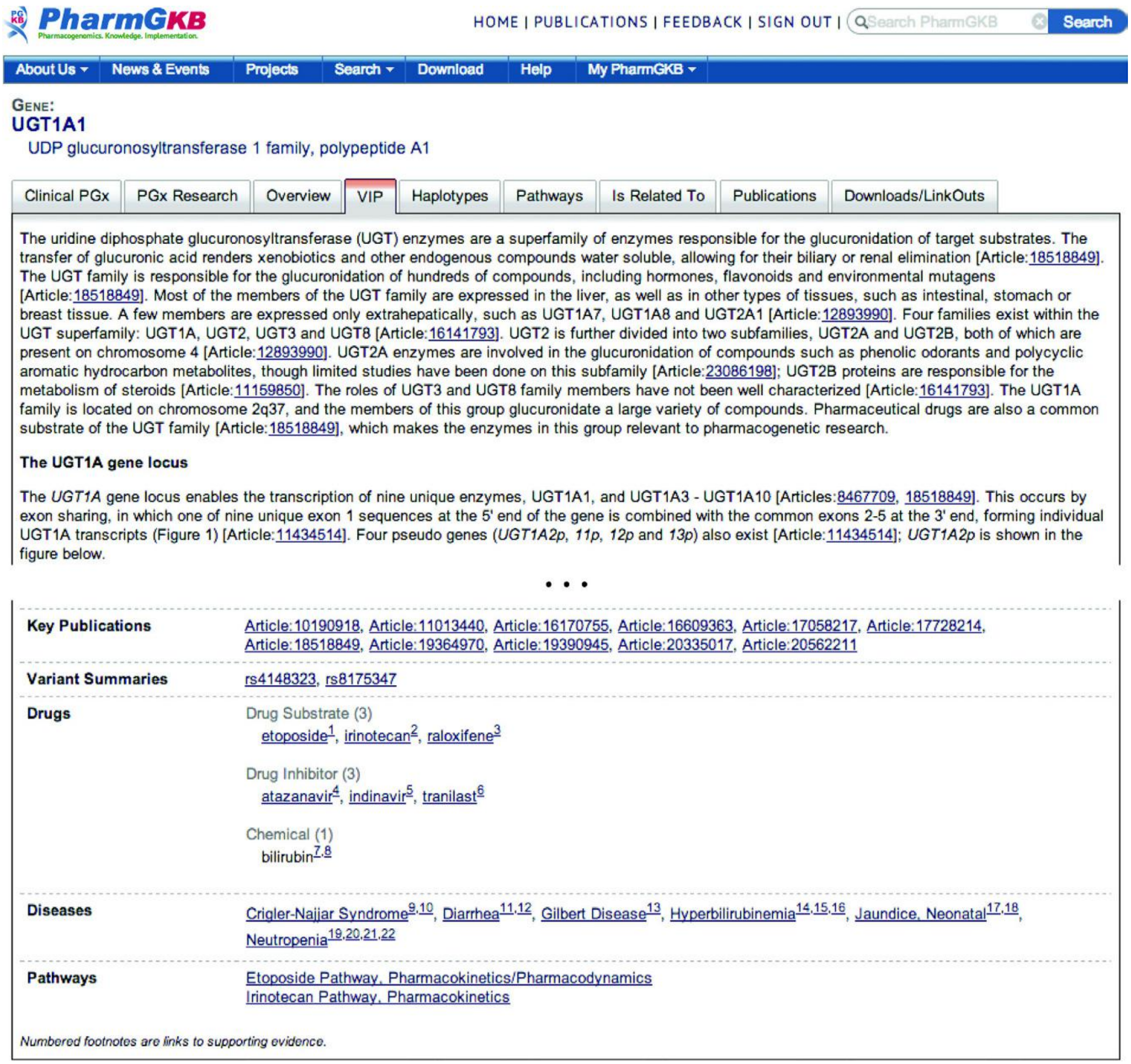
Annotation based on a case report, non-significant study or in vitro, molecular or functional assay evidence only.

FIGURE 14.7 **PharmGKB levels of evidence for clinical annotations.** Assigning levels of evidence to an association is a critical function of PharmGKB curators. Internal guidelines for assigning levels of evidence provide consistency for curators and transparency for users. Many annotations are unique and therefore are assigned a level of 3. However, as annotations accumulate for a particular variant–drug response, the level of evidence may change.

relevant. Associations that show level 1B evidence are of particular importance to PharmGKB, as they likely have potential for clinical implementation; level 1A indicates that implementation is already occurring or that a PGx guideline exists for that variant and drug. The association between rs776746 and tacrolimus dose and metabolism is assigned an evidence level of 1B, and PharmGKB has begun work with the Clinical Pharmacogenentics Implementation Consortium (CPIC) [15] to develop dosing guidelines for the drug based on the rs776746 genotype. PharmGKB keeps a close eye on all associations that show high levels of evidence, and the resource often works closely with consortia on guidelines or implementation for associations with strong supporting evidence. As with variant annotation data, registered users who submit a request can download a file of all PharmGKB clinical annotation data if they wish to pursue further research using this information. Indeed, various implementation research programs, such as the 1200 Patients Project, have used clinical annotations to help decide which PGx variants to include in their studies [16].

14.5 VIPS AND PATHWAYS

In addition to variant and clinical annotations, PharmGKB creates informational materials about genes and drugs of particular pharmacogenomic interest. VIPs are overviews of important genes involved in drug response that typically include background information on the gene and any disease associations, as well as in-depth information on the gene's pharmacogenetics. *CYP3A5*, the gene that contains rs776746, has a VIP; it is part of the *CYP3A* gene family, whose protein members facilitate the elimination of a



GENE:
UGT1A1
UDP glucuronosyltransferase 1 family, polypeptide A1

The uridine diphosphate glucuronosyltransferase (UGT) enzymes are a superfamily of enzymes responsible for the glucuronidation of target substrates. The transfer of glucuronic acid renders xenobiotics and other endogenous compounds water soluble, allowing for their biliary or renal elimination [Article:18518849]. The UGT family is responsible for the glucuronidation of hundreds of compounds, including hormones, flavonoids and environmental mutagens [Article:18518849]. Most of the members of the UGT family are expressed in the liver, as well as in other types of tissues, such as intestinal, stomach or breast tissue. A few members are expressed only extrahepatically, such as UGT1A7, UGT1A8 and UGT2A1 [Article:12893990]. Four families exist within the UGT superfamily: UGT1A, UGT2, UGT3 and UGT8 [Article:16141793]. UGT2 is further divided into two subfamilies, UGT2A and UGT2B, both of which are present on chromosome 4 [Article:12893990]. UGT2A enzymes are involved in the glucuronidation of compounds such as phenolic odorants and polycyclic aromatic hydrocarbon metabolites, though limited studies have been done on this subfamily [Article:23086198]; UGT2B proteins are responsible for the metabolism of steroids [Article:11159850]. The roles of UGT3 and UGT8 family members have not been well characterized [Article:16141793]. The UGT1A family is located on chromosome 2q37, and the members of this group glucuronidate a large variety of compounds. Pharmaceutical drugs are also a common substrate of the UGT family [Article:18518849], which makes the enzymes in this group relevant to pharmacogenetic research.

The UGT1A gene locus

The *UGT1A* gene locus enables the transcription of nine unique enzymes, UGT1A1, and UGT1A3 - UGT1A10 [Articles:8467709, 18518849]. This occurs by exon sharing, in which one of nine unique exon 1 sequences at the 5' end of the gene is combined with the common exons 2-5 at the 3' end, forming individual UGT1A transcripts (Figure 1) [Article:11434514]. Four pseudo genes (*UGT1A2p*, *11p*, *12p* and *13p*) also exist [Article:11434514]; *UGT1A2p* is shown in the figure below.

• • •

Key Publications	Article:10190918, Article:11013440, Article:16170755, Article:16609363, Article:17058217, Article:17728214, Article:18518849, Article:19364970, Article:19390945, Article:20335017, Article:20562211
Variant Summaries	rs4148323, rs8175347
Drugs	Drug Substrate (3) etoposide[1], irinotecan[2], raloxifene[3] Drug Inhibitor (3) atazanavir[4], indinavir[5], tranilast[6] Chemical (1) bilirubin[7,8]
Diseases	Crigler-Najjar Syndrome[9,10], Diarrhea[11,12], Gilbert Disease[13], Hyperbilirubinemia[14,15,16], Jaundice, Neonatal[17,18], Neutropenia[19,20,21,22]
Pathways	Etoposide Pathway, Pharmacokinetics/Pharmacodynamics Irinotecan Pathway, Pharmacokinetics

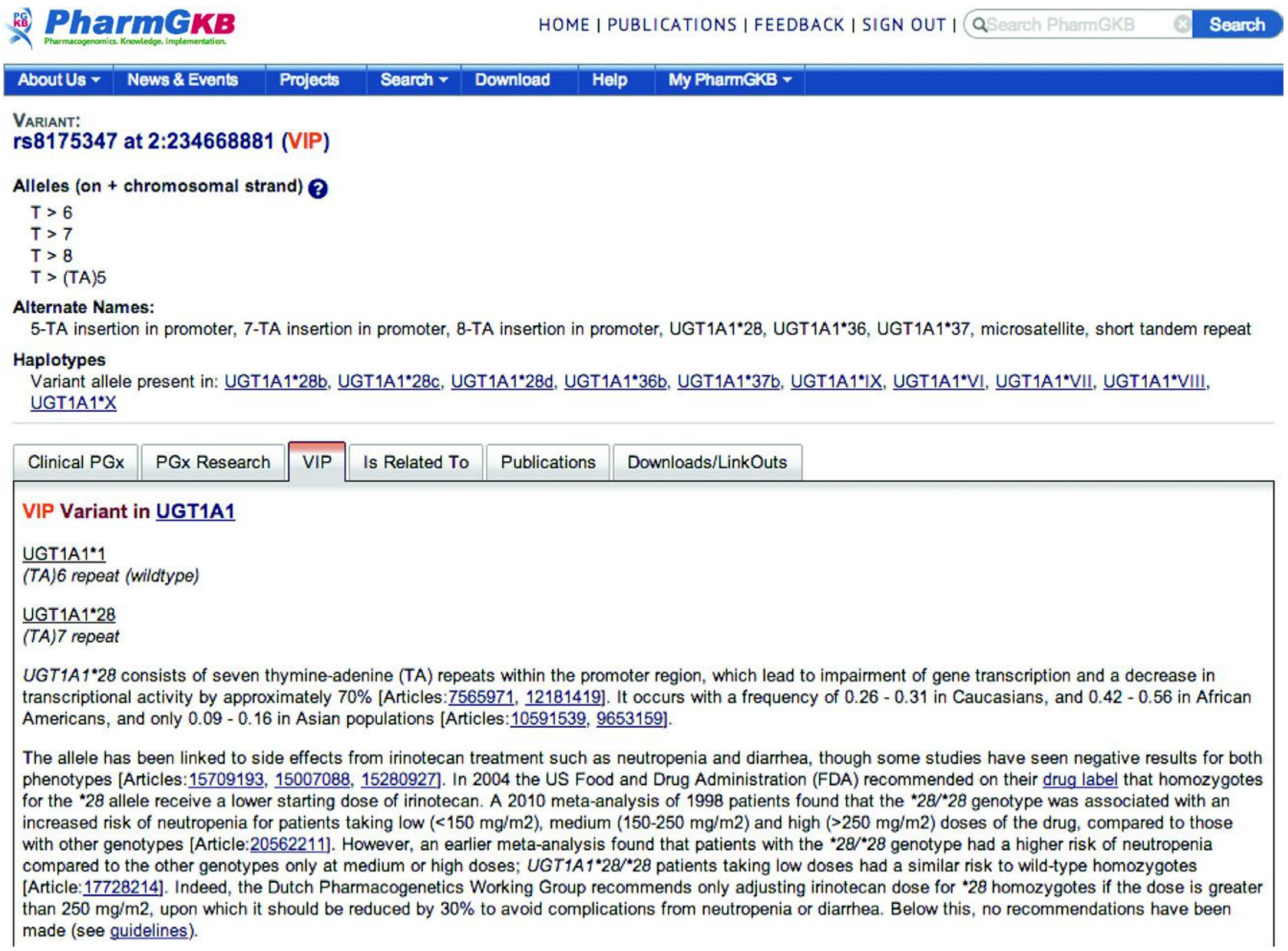
Numbered footnotes are links to supporting evidence.

FIGURE 14.8 **VIP for the gene *UGT1A1*.** VIPs are summaries of genes of particular pharmacogenomic interest. The UGT1A1 protein is part of the UGT family and is involved in the glucuronidation of a wide variety of pharmaceuticals, making it relevant to PGx research. VIPs consist of textual summaries (a portion of which is shown in the figure), as well as lists of key publications and drugs and diseases that are relevant to the gene. Any pathways in which the gene appears are also listed. VIPs, if available, can be accessed through the "VIP" tab at the top of the page.

large percentage of drugs. Other examples of VIPs include *G6PD*, *TMPT*, and *UGT1A1*; a list of all PharmGKB VIPs can be found on the website. Individual VIPs, if available, can be accessed through the relevant PharmGKB gene page.

Figure 14.8 is an example of a VIP for the gene *UGT1A1*. A portion of the textual summary is shown, as is the bottom of the page, where key publications are cited. VIPs are intended for users wishing to gain a better understanding of a gene, and key publications represent references that provide this for the gene and its pharmacogenetics. An overview of the drugs and diseases relevant to the gene are also given at the bottom of the summary, with links to their corresponding PharmGKB pages.

VIPs often include independent summaries for the important PGx variants within a gene. If available, these can be accessed through the "Variant Summaries" row at the bottom of the VIP page (see Figure 14.8). An example of a variant summary for rs8175347 in *UGT1A1* (also known as *UGT1A1**28) is shown in Figure 14.9. VIP variant summaries complement the clinical annotations for a variant by providing more background information, such as the biological effect of the allele on the gene, allele frequencies,



VARIANT:
rs8175347 at 2:234668881 (VIP)

Alleles (on + chromosomal strand)
T > 6
T > 7
T > 8
T > (TA)5

Alternate Names:
5-TA insertion in promoter, 7-TA insertion in promoter, 8-TA insertion in promoter, UGT1A1*28, UGT1A1*36, UGT1A1*37, microsatellite, short tandem repeat

Haplotypes
Variant allele present in: UGT1A1*28b, UGT1A1*28c, UGT1A1*28d, UGT1A1*36b, UGT1A1*37b, UGT1A1*IX, UGT1A1*VI, UGT1A1*VII, UGT1A1*VIII, UGT1A1*X

VIP Variant in UGT1A1

UGT1A1*1
(TA)6 repeat (wildtype)

UGT1A1*28
(TA)7 repeat

*UGT1A1*28* consists of seven thymine-adenine (TA) repeats within the promoter region, which lead to impairment of gene transcription and a decrease in transcriptional activity by approximately 70% [Articles:7565971, 12181419]. It occurs with a frequency of 0.26 - 0.31 in Caucasians, and 0.42 - 0.56 in African Americans, and only 0.09 - 0.16 in Asian populations [Articles:10591539, 9653159].

The allele has been linked to side effects from irinotecan treatment such as neutropenia and diarrhea, though some studies have seen negative results for both phenotypes [Articles:15709193, 15007088, 15280927]. In 2004 the US Food and Drug Administration (FDA) recommended on their drug label that homozygotes for the *28 allele receive a lower starting dose of irinotecan. A 2010 meta-analysis of 1998 patients found that the **28/*28* genotype was associated with an increased risk of neutropenia for patients taking low (<150 mg/m2), medium (150-250 mg/m2) and high (>250 mg/m2) doses of the drug, compared to those with other genotypes [Article:20562211]. However, an earlier meta-analysis found that patients with the **28/*28* genotype had a higher risk of neutropenia compared to the other genotypes only at medium or high doses; *UGT1A1*28/*28* patients taking low doses had a similar risk to wild-type homozygotes [Article:17728214]. Indeed, the Dutch Pharmacogenetics Working Group recommends only adjusting irinotecan dose for *28 homozygotes if the dose is greater than 250 mg/m2, upon which it should be reduced by 30% to avoid complications from neutropenia or diarrhea. Below this, no recommendations have been made (see guidelines).

FIGURE 14.9 **VIP Variant summary for rs8175347.** VIP variant summaries are for variants that show important PGx associations; they are often created in tandem with VIPs. If available, these overviews can be accessed through the "VIP" tab on a variant page. They provide information about the variant, including allele frequencies and a detailed overview of the literature evidence for any drug–response associations.

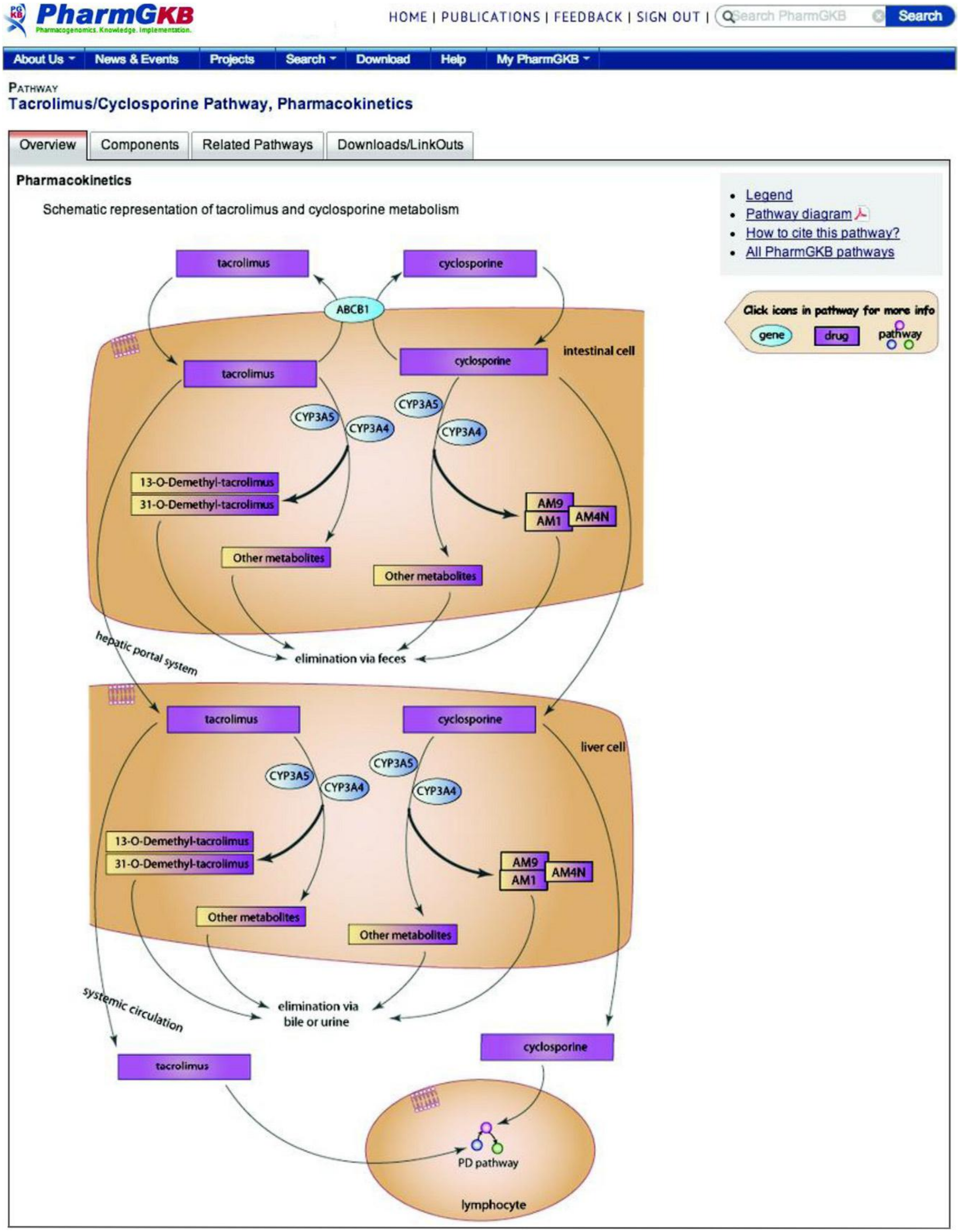
and a more detailed overview of the literature evidence for that variant. Variant summaries can be found on the relevant variant's page.

VIP genes are chosen through a variety of means. These genes include those that are referred to on U.S. Food and Drug Administration (FDA) drug labels or that appear as part of important pharmacogenomic developments discussed in recent literature. Additionally, VIPs can be created if a curator notes that the gene is associated with a large number of variant annotations or is part of high-level clinical annotations.

PharmGKB pathways are illustrations detailing the pharmacokinetics or pharmacodynamics of a PGx-relevant drug; they demonstrate physiological interactions between the drug and various proteins (represented as the corresponding genes in PharmGKB) in cells. Figure 14.10 shows the pharmacokinetic pathway created for the metabolism of tacrolimus and cyclosporine. These two drugs, both immunosuppressants, have identical pharmacokinetics and so appear together in the pathway.

Other pathways on PharmGKB include warfarin pharmacokinetics and pharmacodynamics, and tamoxifen pharmacokinetics. A list of all PharmGKB pathways can be found on the website. PharmGKB recognizes that proteins are responsible for metabolizing or transporting drugs; however, these products are stored as genes in the knowledgebase, and so the genes coding for various proteins are used in pathways. In Figure 14.10, *CYP3A5* (coding for its eponymous enzyme) is shown metabolizing tacrolimus, while *ABCB1* (coding for P-glycoprotein) is shown transporting the drug across the intestinal cell membrane. PharmGKB pathways are interactive—drugs and genes that appear on them are clickable and link to their respective pages. Pathways can thus be accessed from the relevant drug or gene pages. Additionally, pathways are downloadable in multiple formats, including BioPax, to allow the information stored in them to be easily used in different ways.

PharmGKB pathways are manually created by curators after an extensive literature review, and the connections on the pathway in the online version of the figure are supported


PharmGKB
Pharmacogenomics. Knowledge. Implementation.
HOME | PUBLICATIONS | FEEDBACK | SIGN OUT |
Search PharmGKB
Search
About Us
News & Events
Projects
Search
Download
Help
My PharmGKB
Overview
Components
Related Pathways
Downloads/LinkOuts

PATHWAY

Tacrolimus/Cyclosporine Pathway, Pharmacokinetics

Pharmacokinetics

Schematic representation of tacrolimus and cyclosporine metabolism

- Legend
- Pathway diagram
- How to cite this pathway?
- All PharmGKB pathways

Click icons in pathway for more info

gene
drug
pathway
tacrolimus
cyclosporine
ABCB1
intestinal cell
tacrolimus
cyclosporine
CYP3A5
CYP3A4
CYP3A5
CYP3A4
13-O-Demethyl-tacrolimus
31-O-Demethyl-tacrolimus
AM9
AM1
AM4N
Other metabolites
Other metabolites
hepatic portal system
elimination via feces
liver cell
tacrolimus
cyclosporine
CYP3A5
CYP3A4
CYP3A5
CYP3A4
13-O-Demethyl-tacrolimus
31-O-Demethyl-tacrolimus
AM9
AM1
AM4N
Other metabolites
Other metabolites
systemic circulation
elimination via bile or urine
tacrolimus
cyclosporine
PD pathway
lymphocyte

FIGURE 14.10 **PharmGKB tacrolimus and cyclosporine pharmacokinetics pathway.** Pathways detail either the pharmacokinetics or pharmacodynamics of a drug of interest. Genes that code for proteins involved in a drug's pharmacology are shown on the pathway, which always includes accompanying text that provides more information about the pharmacology and pharmacogenomics of the drug. They are excellent starting points for learning more about a PGx-relevant medication.

by specific literature citations; these can be found under the "Components" tab on the pathway page (see Figure 14.10). Pathway images are accompanied by text that provides more general information about the drug, as well as a discussion of its pharmacokinetics or pharmacodynamics and pharmacogenomics. Pathways are an excellent starting point for users who want to learn more about a drug, and PharmGKB pathways are often used in lectures and book chapters. Pathways are created for drugs that have enough literature-based evidence to support graphical, gene-based representations of the pharmacology of a drug. Both pathways and VIPs are often created in collaboration with experts in the field and are typically published in the journal *Pharmacogenetics and Genomics*.

14.6 CLINICAL IMPLEMENTATION: CONSORTIA, CPIC, AND DOSING GUIDELINES

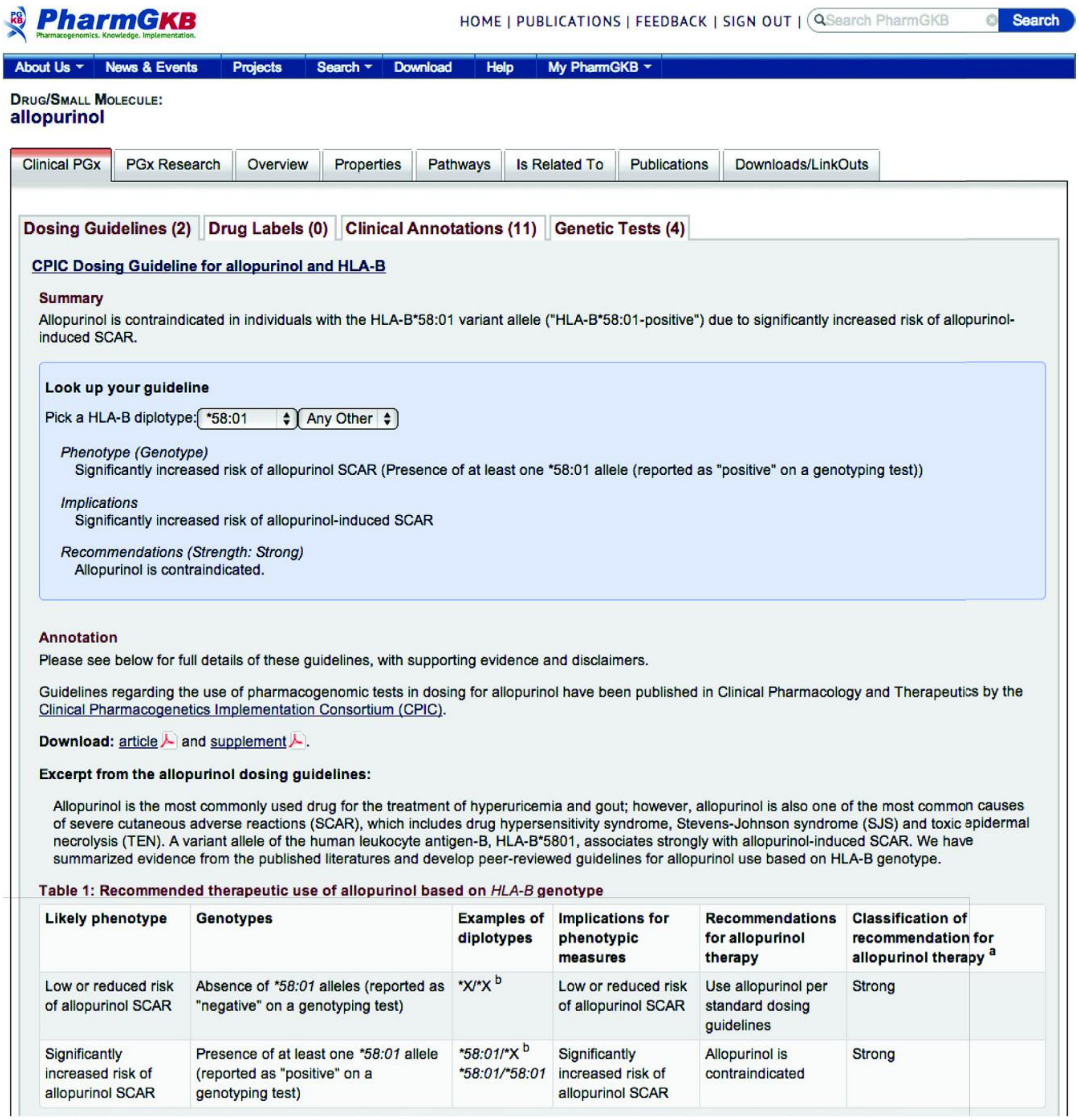
In addition to storing PGx knowledge and creating unique materials, PharmGKB also works closely with various consortia toward a common goal of clinical implementation. It is involved with two basic kinds of consortia: data-centric and knowledge-centric. Data-centric consortia consist of domestic and international groups conducting similar PGx studies. The goal is to gather their various study results into one or more analyses, increasing available ethnicities as well as study size and statistical power. Through these consortia, complicated genetic questions can be clarified and new and important pharmacogenomic variants may be discovered. The International Warfarin Pharmacogenetics Consortium (IWPC) is responsible for collating datasets on warfarin dosing from various pharmacogenomic research groups, and has so far released several papers on its findings [17,18] as well as a global dosing algorithm that utilizes the *VKORC1* and *CYP2C9* genotypes [19].

While dosing algorithms for warfarin had been previously proposed, they were usually based on data from small clinical populations and so had uncertain predictive accuracy. By bringing together data from over 4000 patients from 21 research groups around the world, the IWPC has been able to propose a reliable pharmacogenetic dosing algorithm that provides significantly more accurate dose estimates than a clinical dosing algorithm or a fixed-dose approach [19]. Because of its global reach, the IWPC analysis cohort consists of large numbers of black and Asian participants in addition to whites [19]. *VKORC1* polymorphisms are known to account for a smaller percentage of dose variability in blacks compared to whites or Asians, but prior to the formation of the IWPC, research on this ethnicity was difficult owing to low numbers of black participants in study cohorts [18]. The multiethnic nature of IWPC participants allowed the consortium to assess whether incorporating additional *VKORC1* variants into the dosing algorithm improved dose prediction in blacks or Asians [18], and perform a genome-wide association study to identify additional variants that may contribute to dose requirements in blacks [17].

Other consortia include the International Clopidogrel Pharmacogenomics Consortium (ICPC) and the International Consortium for Antihypertensives Pharmacogenomics Studies (ICAPS), which work to discover the genetic causes of variable response to clopidogrel and antihypertensives, respectively. PharmGKB acts as an independent broker of information in these consortia and therefore plays a crucial role in ensuring their success. It has no vested interest in the outcome and so can facilitate collaboration and analyses of these data. PharmGKB curators are experienced at harmonizing data across phenotypes, genotypes, and cultural differences, and they are therefore invaluable in keeping vast quantities of data organized. Current consortia are described on the PharmGKB website (www.pharmgkb.org/page/projects), and the knowledgebase is always open to new opportunities.

CPIC is a knowledge-centric consortium, consisting of members of NIH's Pharmacogenomics Research Network (www.pgrn.org), PharmGKB staff, and experts in both pharmacogenomics and medicine. While the goal of consortia such as IWPC, ICPC, and ICAPS is to amass datasets to answer a research objective, the aim of CPIC is to create detailed drug-dosing guidelines for clinicians who have access to pre-emptive genetic test results [15]. One of the major barriers to clinical implementation is a lack of freely available, peer-reviewed, updatable guidelines, and CPIC works to address this deficiency. Its guidelines target PGx associations that show substantial positive evidence, such as *HLA-B*58:01* and allopurinol [20] or *SLCO1B1* rs4149056 and simvastatin [21]. PharmGKB may suggest a CPIC guideline for an association based on a high-evidence-level clinical annotation.

Guidelines are written in a standardized format, consisting of background information on the gene, variant, and drug being discussed, information on genetic test options, and a discussion of the evidence linking genotype to phenotype. They culminate in specific therapeutic recommendations written in text form and presented as a table. Each guideline is written after an extensive literature review; PharmGKB staff and other CPIC members work in collaboration to conduct this review as well as to write the manuscript. Each therapeutic recommendation is given a strength rating, based on the quality and quantity of evidence as well as on whether the benefits of following the recommendations outweigh any possible undesirable effects. Most dosing recommendations receive a rating of "Strong." Ratings allow clinicians to make better-informed decisions about possible dosing modifications. More on these rating schemes can be found in the review by Relling and Klein [15] and in the supplementary material for any CPIC guideline.



DRUG/SMALL MOLECULE:
allopurinol

CPIC Dosing Guideline for allopurinol and HLA-B

Summary

Allopurinol is contraindicated in individuals with the HLA-B*58:01 variant allele ("HLA-B*58:01-positive") due to significantly increased risk of allopurinol-induced SCAR.

Look up your guideline

Pick a HLA-B diplotype: *58:01 | Any Other

Phenotype (Genotype)
Significantly increased risk of allopurinol SCAR (Presence of at least one *58:01 allele (reported as "positive" on a genotyping test))

Implications
Significantly increased risk of allopurinol-induced SCAR

Recommendations (Strength: Strong)
Allopurinol is contraindicated.

Annotation

Please see below for full details of these guidelines, with supporting evidence and disclaimers.

Guidelines regarding the use of pharmacogenomic tests in dosing for allopurinol have been published in Clinical Pharmacology and Therapeutics by the Clinical Pharmacogenetics Implementation Consortium (CPIC).

Download: article and supplement.

Excerpt from the allopurinol dosing guidelines:

Allopurinol is the most commonly used drug for the treatment of hyperuricemia and gout; however, allopurinol is also one of the most common causes of severe cutaneous adverse reactions (SCAR), which includes drug hypersensitivity syndrome, Stevens-Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN). A variant allele of the human leukocyte antigen-B, HLA-B*5801, associates strongly with allopurinol-induced SCAR. We have summarized evidence from the published literatures and develop peer-reviewed guidelines for allopurinol use based on HLA-B genotype.

Table 1: Recommended therapeutic use of allopurinol based on *HLA-B* genotype

Likely phenotype	Genotypes	Examples of diplotypes	Implications for phenotypic measures	Recommendations for allopurinol therapy	Classification of recommendation for allopurinol therapy [a]
Low or reduced risk of allopurinol SCAR	Absence of **58:01* alleles (reported as "negative" on a genotyping test)	*X/*X [b]	Low or reduced risk of allopurinol SCAR	Use allopurinol per standard dosing guidelines	Strong
Significantly increased risk of allopurinol SCAR	Presence of at least one **58:01* allele (reported as "positive" on a genotyping test)	**58:01/*X* [b] **58:01/*58:01*	Significantly increased risk of allopurinol SCAR	Allopurinol is contraindicated	Strong

FIGURE 14.11 **CPIC guideline for *HLA-B**58:01 and allopurinol.** In addition to being published in *Clinical Pharmacology and Therapeutics*, CPIC guidelines are distributed through PharmGKB. Dosing guidelines can be found under the "Clinical PGx" and "Dosing Guidelines" tabs on the relevant gene and drug pages. Each page includes an excerpt from the guideline, the therapeutic dosing recommendations in table format, and links to the original article and supplement (the supplement includes information about the strength-rating scheme). PharmGKB also includes interactive dosing guidelines, which can be seen at the top of the page. Users can enter a genotype of interest and receive specific dosing guidelines for it taken from the dosing recommendations table.

CPIC guidelines are published in *Clinical Pharmacology and Therapeutics* and are open-source on PubMed Central. PharmGKB provides a list of published guidelines with links to PubMed at www.pharmgkb.org/page/cpicGeneDrugPairs; these guidelines are available for download from the website in JSON format. PharmGKB also presents the guidelines on its website in a simplified format, accessible through the relevant gene or drug pages. An example of the CPIC guideline for *HLA-B**58:01 and allopurinol is shown in Figure 14.11. An excerpt is included, as well as links to the original paper and the therapeutic dosing guideline table. CPIC guidelines on PharmGKB also include interactive dosing guidelines, where users can enter in their specific genotype of interest and receive dosing information sourced from the therapeutic dosing recommendations table. This functionality can be helpful for guidelines that have very large and extensive recommendation tables, or for users looking for quick information. Figure 14.11 shows the response to a request for information on *HLA-B**58:01 heterozygotes: for these patients, allopurinol is contraindicated because of an increased risk of severe cutaneous adverse reactions (SCARs).

In addition to distributing CPIC guidelines through its website, PharmGKB also sources dosing guidelines from other groups, such as The Royal Dutch Association for the Advancement of Pharmacy Pharmacogenetics Working Group (DWPG) [22]. A list of all dosing guidelines on PharmGKB can be found at www.pharmgkb.org/view/dosing-guidelines.do; guidelines from CPIC and DWPG are added as they are released. PharmGKB also sources drug labels that contain PGx information from the European Medicines Agency (EMA) [23] and from FDA-approved drug labels with PGx information from DailyMed [24,25]. If available, these can be found on their relevant gene and drug pages. PharmGKB presents an excerpt from the drug label as well as a downloadable copy of the label with the PGx information highlighted.

14.7 PROJECTS IN DEVELOPMENT

Data on PGx literature form the basis of all materials created on PharmGKB, and a current challenge for the knowledgebase is keeping pace with the large volume of literature being published. PharmGKB is integrating NLP technology to assist curators in the identification and extraction of PGx relationships from the literature, through a pipeline known as SemanticPGx, which automatically extracts novel gene–drug associations from PubMed. The pipeline begins with a download of all PubMed abstracts and, after several steps, including parsing and filtration, presents a collection of novel relationships ranked by the number of PubMed abstracts in which they appear. Novel relationships are defined as those between a gene and a drug that are not currently present on PharmGKB, where the gene in the relationship has a relationship with other PharmGKB drugs and the drug has a relationship with other PharmGKB genes. Curators can view this assortment of extracted relationships, assess them for relevancy and accuracy, and then consider annotation of the collected publications. This allows curators to maintain accuracy while permitting broader literature coverage and greater data input.

In addition to expanding the knowledgebase, PharmGKB is working on making the currently available data more easily accessible to both researchers and clinicians. This is achieved through two research projects: PGxplore and PGxpress. PGxplore is an interface that allows visualization of all PGx relationships in the knowledgebase. When entities are listed together in a variant annotation, clinical annotation, VIP summary, pathway, dosing guideline, or drug label, they are stored in the knowledgebase as a relationship. This creates a vast network of connections from multiple types of data. PGxplore allows users to visualize these links in a force-directed and entity-centered manner. Figure 14.12A shows a PGxplore image created for the drug tacrolimus, although complicated networks centered on drugs, genes, diseases, phenotypes, or variants can be created. Different-color lines represent the different sources from which the relationship is derived, showing the way in which tacrolimus is connected with a variety of genes and diseases through different PharmGKB sources.

As an example, Figure 14.12B zooms in on the connection between tacrolimus and *CYP3A5*. We have shown how these two entities appear together in a variant annotation (refer to Figure 14.3), a clinical annotation (refer to Figure 14.6), and a pathway (refer to Figure 14.10). They also appear together in a VIP as well as in a dosing guideline. In the image, the purple line shows that tacrolimus and *CYP3A5* are connected through a variant annotation; the blue line indicates a relationship through a clinical annotation; all other lines indicate relationships through other materials. A user can click on any connection to see supporting data.

This type of visualization provides an alternate way to learn about the interaction of genes, drugs, and phenotypes on PharmGKB, as compared to scrolling through tables. The networks can be complex, but the PGxplore environment provides many ways to simplify them (or make them more complex if that is desired), and a number of tools and options are available for exploring the data in depth. Indeed, data from any outside source can be uploaded to the platform as long as they are in an accepted format—PGxplore accepts data in many forms, including JSON, XGMML, OWL, and CSV, which allows for an expansion of research opportunities beyond PharmGKB. Overall, PGxplore provides an interactive and visually agreeable way to manipulate and comprehend data from PharmGKB or from outside sources, and it acts as a platform for the discovery of novel PGx associations and subsequent hypothesis generation.

While PGxplore focuses on visualizing domain knowledge and increasing accessibility to PharmGKB data, PGxpress streamlines PharmGKB knowledge for clinicians or researchers. PGxpress is a PGx mobile website that targets users who want high-impact PharmGKB data on smart phones or tablets. The main PharmGKB website is designed to be displayed on desktop screens, and pages often display extensive text and tables that are ideal for viewing large amounts of detailed content. However, navigating the website from a phone or tablet can be difficult for a researcher looking up a variant of clinical interest while away from his or her desk, or for a clinician with a patient wanting to check dosing guidelines. Gene and drug pages on PGxpress present clinically relevant information in condensed form, including dosing guidelines from CPIC and DWPG, drug labels, and annotation data. An example of the condensed

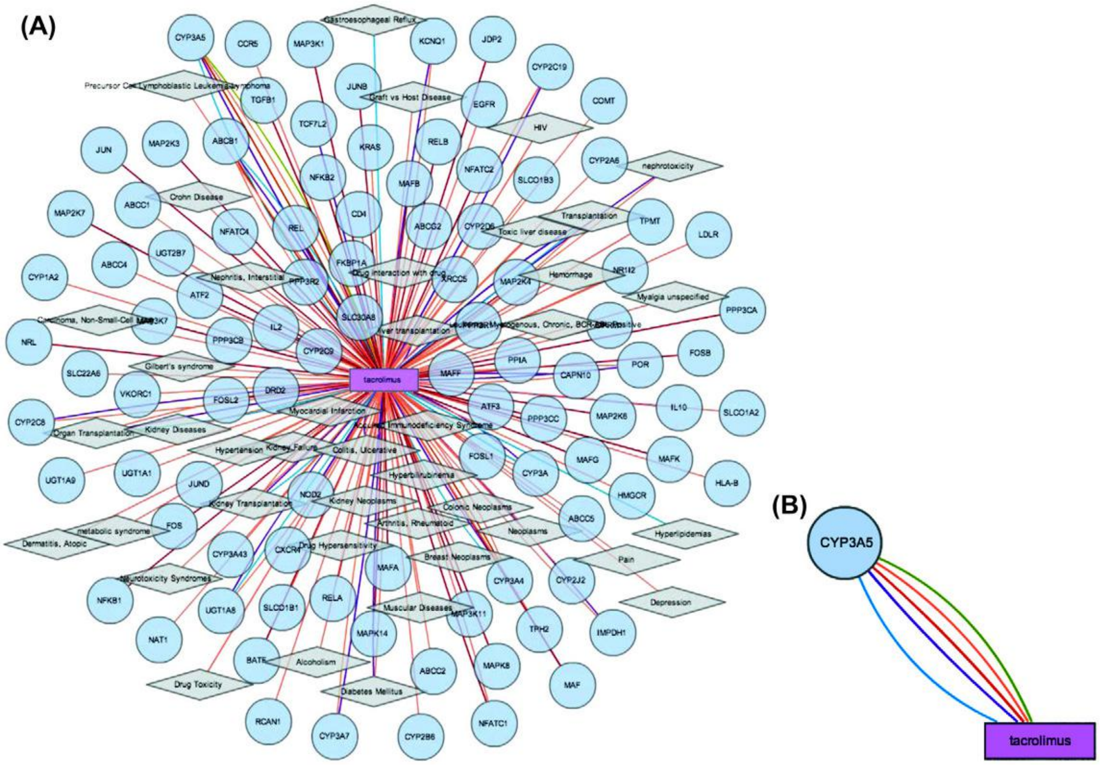
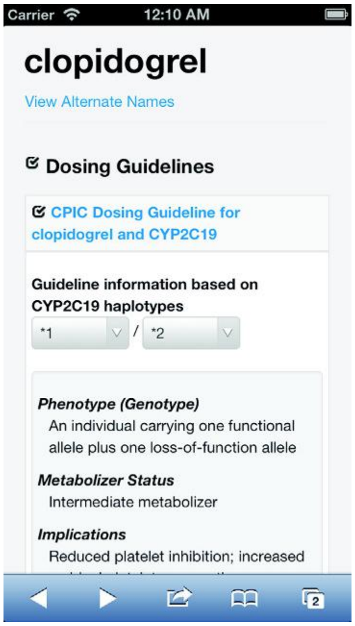

(A)
CYP3A5
CCR5
MAP3K1
Gastroesophageal Reflux
KCNQ1
JDP2
CYP2C19
Precursor Cell Lymphoblastic Leukemia-Lymphoma
TGFB1
JUNB
Graft vs Host Disease
EGFR
COMT
TCF7L2
HIV
JUN
MAP2K3
ABCB1
KRAS
RELB
NFATC2
CYP2A6
nephrotoxicity
NFKB2
MAFB
SLCO1B3
Crohn Disease
MAP2K7
ABCC1
NFATC4
REL
CD4
ABCG2
CYP2D6
Transplantation
TPMT
LDLR
Toxic liver disease
UGT2B7
CYP1A2
ABCC4
FKBP1A
Drug interaction with drug
Nephritis, Interstitial
PPP3R2
ERCC5
MAP2K4
Hemorrhage
NR1I2
ATF2
Myalgia unspecified
PPP3CA
Carcinoma, Non-Small-Cell Lung
MAP3K7
SLC30A8
IL2
liver transplantation
Leukemia, Myelogenous, Chronic, BCR-ABL Positive
NRL
PPP3CB
CYP2C9
PPIA
FOSB
Gilbert's syndrome
SLC22A6
tacrolimus
MAFF
CAPN10
POR
VKORC1
DRD2
FOSL2
Myocardial Infarction
ATF3
IL10
SLCO1A2
CYP2C8
PPP3CC
MAP2K6
Organ Transplantation
Kidney Diseases
Acquired Immunodeficiency Syndrome
Hypertension
Kidney Failure
Colitis, Ulcerative
FOSL1
MAFG
MAFK
UGT1A9
UGT1A1
JUND
CYP3A
HLA-B
Hyperbilirubinemia
NOD2
Kidney Neoplasms
HMGCR
Kidney Transplantation
Colonic Neoplasms
Arthritis, Rheumatoid
metabolic syndrome
FOS
Neoplasms
ABCC5
Hyperlipidemias
Dermatitis, Atopic
CYP3A43
CXCR4
Drug Hypersensitivity
Breast Neoplasms
Pain
Neurotoxicity Syndromes
MAFA
CYP3A4
CYP2J2
NFKB1
UGT1A8
SLCO1B1
RELA
Depression
Muscular Diseases
MAP3K11
NAT1
MAPK14
TPH2
IMPDH1
BATF
Alcoholism
ABCC2
MAPK8
Drug Toxicity
MAF
Diabetes Mellitus
RCAN1
NFATC1
CYP3A7
CYP2B6
(B)
CYP3A5
tacrolimus

FIGURE 14.12 PGxplore images for tacrolimus. **(A)** A drug-centered complicated network created for tacrolimus. The lines between tacrolimus and the various genes or diseases indicate that these two entities appear together in one or more PharmGKB materials. PGxplore provides multiple ways to simplify networks, if necessary, or to increase their complexity if desired. **(B)** Connection between tacrolimus and *CYP3A5* only. Different-color lines indicate the different sources from which the connection is derived. For instance, the purple line indicates that the two entities appear together in at least one variant annotation, while the green line shows that they are used in a dosing guideline.


Carrier
12:10 AM

clopidogrel

View Alternate Names

Dosing Guidelines

CPIC Dosing Guideline for clopidogrel and CYP2C19

Guideline information based on CYP2C19 haplotypes

*1 / *2

Phenotype (Genotype)
An individual carrying one functional allele plus one loss-of-function allele

Metabolizer Status
Intermediate metabolizer

Implications
Reduced platelet inhibition; increased

FIGURE 14.13 PGxpress page for clopidogrel. PGxpress presents clinically relevant information in a condensed format that is ideal for any screen size. From gene, drug, or variant pages, users can check for clinically relevant information, including dosing guidelines, drug labels, and annotation data.

page for clopidogrel, as viewed on an iPhone, is shown in Figure 14.13.

In addition, users can easily check for any necessary dosing modifications using drop-down boxes, and the overall format of PGxpress means this information is legible on any size screen. Links are also provided to the original drug, gene, or variant page on PharmGKB if more information is needed. Through PGxpress, PharmGKB provides direct access to key PGx information in a mobile format, expanding its data beyond the desktop and into the pockets of clinicians and researchers.

14.8 CONCLUSION

Through its collection of literature and data, PharmGKB has grown into the foremost knowledgebase of PGx information. Moreover, by providing this knowledge in multiple formats, it makes its information accessible and understandable to all types of users. Whereas researchers may choose to use variant annotation data for their studies, students may find VIPs and pathways ideal for learning about pharmacogenomics. Those interested in clinically relevant information have access to curated and evidence-rated clinical annotations. Through the efforts of CPIC and other consortia, PharmGKB can provide dosing guidelines for interested clinicians and help advance clinical implementation. Given the magnitude of knowledge available on its website, PharmGKB has worked carefully to ensure that all available information is highly curated, well organized, and easy to find. In addition to growing the knowledgebase, PharmGKB continues to work on new projects that increase the usability of its knowledge for both researchers and clinicians. Through these undertakings, PharmGKB endeavors to both broaden its user base and further pharmacogenomic knowledge.

Key Points about PharmGKB

What is the Pharmacogenomics Knowledgebase (PharmGKB)?
PharmGKB is a publicly available, online resource that provides information about the effect of human genetic variations on drug response. It also plays a large role in various consortia, which work globally to answer important pharmacogenomic questions and implement pharmacogenomics in the clinic.

What types of information can I find on PharmGKB?
PharmGKB extracts information from primary pharmacogenomic literature and presents it in the form of variant annotations; these data are also available for download for further research. The knowledgebase aggregates and integrates its data to create higher-level materials such as clinical annotations, very important pharmacogene (VIP) summaries, and pathways. Additionally, the website distributes information gathered from its work with various consortia, such as the drug-dosing guidelines developed as part of the Clinical Pharmacogenetics Implementation Consortium (CPIC). It also curates drug-dosing guidelines from other sources as well as U.S. and European drug labels that contain pharmacogenomic information.

What is a variant annotation?
A variant annotation is a summary of an association between a single variant and a drug phenotype as reported in a publication. A publication can have as many annotations as different associations found.

What is a clinical annotation?
A clinical annotation is an aggregation of variant annotations. It is a clinically oriented and genotype-based summary of a particular variant–drug phenotype association. Each clinical annotation is assigned a level of evidence based on the perceived strength of the association; these evidence levels provide a measure of the potential clinical relevance of a variant–drug phenotype association.

What are VIPs and pathways?
VIPs are overviews of important genes involved in drug response. Pathways are illustrations of the pharmacokinetics or pharmacodynamics of a pharmacogenomics-relevant drug, accompanied by informative text. Both VIPs and pathways are excellent starting places for a user to learn more about a pharmacogenomically important gene or drug.

What is PharmGKB's involvement in consortia?
PharmGKB plays a major role in various consortia that help drive clinical implementation forward. Data-centric consortia, such as the International Warfarin Pharmacogenetics Consortium (IWPC) and the International Clopidogrel Pharmacogenomics Consortium (ICPC), bring together domestic and international groups to address important genetic questions and discover new pharmacogenomic variants. PharmGKB works as an independent broker of information in these consortia and therefore plays an invaluable role in their success. PharmGKB is also involved in knowledge-centric consortia, such as the Clinical Pharmacogenetics Implementation Consortium (CPIC), which develops dosing guidelines for variant-drug associations with high levels of supporting evidence.

What is a CPIC dosing guideline?
CPIC dosing guidelines are recommendations for clinicians who have access to pre-emptive genetic test results; they are written by PharmGKB staff in collaboration with other CPIC members. Guidelines exist for variant–drug associations that show a preponderance of supporting evidence. In addition to background information, they provide specific therapeutic recommendations for a drug based on genotype accompanied by a strength rating. These guidelines are published, and PharmGKB provides interactive versions on its website.

GLOSSARY

Clinical annotation a genotype-based, evidence-rated summary of PharmGKB's literature evidence for a particular variant and drug phenotype association

Pathway an illustration of the pharmacokinetics or pharmacodynamics of a PGx-relevant drug, accompanied by informative text

Pharmacogene a gene involved in drug response

rsID (reference SNP identification) an identification number assigned by the National Center for Biotechnology Information (NCBI) SNP database (dbSNP) to a cluster of SNPs that map to the same location

Variant annotation a summary of an association between a single variant and a drug phenotype as reported in a publication

VIP (very important pharmacogene) summary a pharmacogenetic-based overview of an important gene involved in drug response

ACRONYMS AND ABBREVIATIONS

CPIC	Clinical Pharmacogenetics Implementation Consortium
dbSNP	National Center for Biotechnology Information (NCBI) SNP database
DWPG	The Royal Dutch Association for the Advancement of Pharmacy Pharmacogenetics Working Group
EMA	European Medicines Agency
FDA	U.S. Food and Drug Administration
ICAPS	International Consortium for Antihypertensives Pharmacogenomics Studies
ICPC	International Clopidogrel Pharmacogenomics Consortium
IWPC	International Warfarin Pharmacogenetics Consortium
NLP	natural-language processing
PGx	pharmacogenetics or pharmacogenomics
PharmGKB	Pharmacogenomics Knowledgebase
PMID	PubMed ID
rsID	Reference SNP identification
VIP	very important pharmacogene

ACKNOWLEDGMENTS

PharmGKB is supported by NIH/NIGMS R24 GM61374 and managed at Stanford University.

REFERENCES

[1] Kuehl P, Zhang J, Lin Y, Lamba J, Assem M, Schuetz J, et al. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Genet 2001;27(4):383–91.

[2] Kamdem LK, Streit F, Zanger UM, Brockmoller J, Oellerich M, Armstrong VW, et al. Contribution of CYP3A5 to the *in vitro* hepatic clearance of tacrolimus. Clin Chem 2005;51(8):1374–81.

[3] Niioka T, Satoh S, Kagaya H, Numakura K, Inoue T, Saito M, et al. Comparison of pharmacokinetics and pharmacogenetics of once- and twice-daily tacrolimus in the early stage after renal transplantation. Transplantation 2012;94(10):1013–9.

[4] McDonagh EM, Whirl-Carrillo M, Garten Y, Altman RB, Klein TE. From pharmacogenomic knowledge acquisition to clinical applications: the PharmGKB as a clinical pharmacogenomic biomarker resource. Biomark Med 2011;5(6):795–806.

[5] National Center for Biotechnology Information (NCBI) SNP database (dbSNP). Available at: <www.ncbi.nlm.nih.gov/snp>.

[6] Human Cytochrome P450 (CYP) allele nomenclature database. Available at: <www.cypalleles.ki.se/>.

[7] The UDP-glucuronosyltransferase alleles nomenclature page. Available at: <www.pharmacogenomics.pha.ulaval.ca/cms/ugt_alleles/>.

[8] The HLA nomenclature resource. Available at: <hla.alleles.org>.

[9] The HUGO Gene Nomenclature Committee (HGNC). Available at: www.genenames.org/.

[10] DrugBank. Available at: <www.drugbank.ca>.

[11] WHO Anatomical Therapeutic Chemical (ATC) classification system. Available at: www.whocc.no/atc/structure_and_principles/.

[12] PubChem. Available at: http://pubchem.ncbi.nlm.nih.gov.

[13] National Library of Medicine's Medical Subject Headings (MeSH browser). Available at: <www.nlm.nih.gov/mesh/MBrowser.html>.

[14] Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. Clin Pharmacol Ther 2012;92(4):414–7.

[15] Relling MV, Klein TE. CPIC: Clinical Pharmacogenetics Implementation Consortium of the pharmacogenomics research network. Clin Pharmacol Ther 2011;89(3):464–7.

[16] O'Donnell PH, Bush A, Spitz J, Danahey K, Saner D, Das S, et al. The 1200 patients project: creating a new medical model system for clinical implementation of pharmacogenomics. Clin Pharmacol Ther 2012;92(4):446–9.

[17] Perera MA, Cavallari LH, Limdi NA, Gamazon ER, Konkashbaev A, Daneshjou R, et al. Genetic variants associated with warfarin dose in African-American individuals: a genome-wide association study. Lancet 2013;382(9894):790–6.

[18] Limdi NA, Wadelius M, Cavallari L, Eriksson N, Crawford DC, Lee MT, et al. International warfarin pharmacogenetics consortium.Warfarin pharmacogenetics: a single VKORC1 polymorphism is predictive of dose across 3 racial groups. Blood 2010;115(18):3827–34.

[19] International Warfarin Pharmacogenetics Consortium, Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med 2009;360(8):753–64.

[20] Hershfield MS, Callaghan JT, Tassaneeyakul W, Mushiroda T, Thorn CF, Klein TE, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for human leukocyte antigen-B genotype and allopurinol dosing. Clin Pharmacol Ther 2013;93(2):153–8.

[21] Wilke RA, Ramsey LB, Johnson SG, Maxwell WD, McLeod HL, Voora D, et al. Clinical Pharmacogenomics Implementation Consortium (CPIC). The clinical pharmacogenomics implementation consortium: CPIC guideline for SLCO1B1 and simvastatin-induced myopathy. Clin Pharmacol Ther 2012;92(1):112–7.

[22] Swen JJ, Nijenhuis M, de Boer A, Grandia L, Maitland-van der Zee AH, Mulder H. Pharmacogenetics: from bench to byte--an update of guidelines. Clin Pharmacol Ther 2011;89(5):662–73.

[23] European Medicines Agency (EMA) European Public Assessment Reports (EPAR). Available at: <www.ema.europa.eu/ema/>.

[24] U.S. Food and Drug Administration (FDA) table of pharmacogenomic biomarkers in drug labels. Available at: <www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm>.

[25] DailyMed. Available at: <http://dailymed.nlm.nih.gov/>.

Part V

Clinical Trials

This page is intentionally left blank

Clinical Trials in Pharmacogenomics and Stratified Medicine

Sandosh Padmanabhan
Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Chapter Outline

15.1 Overview 309
15.2 Phase I Studies 309
15.3 Phase II Studies 312
15.4 Phase III Studies 314
  15.4.1 Randomized Controlled Trials 316
15.5 Phase IV Trials 319
15.6 Retrospective Analysis of Clinical Datasets 319
References 319

15.1 OVERVIEW

Drug development is the process of bringing a new drug to the market and consists of multiple stages starting with the discovery of the lead compound, followed by preclinical research (conducted in microorganisms and animals), clinical trials (in humans), and regulatory approval to market the drug (Figure 15.1) [1]. The clinical development of a new drug is fraught with risk and uncertainty, with nearly 90% of human trials failing to achieve registration [2]. One strategy to improve the efficiency and effectiveness of drug development is to leverage pharmacogenomic information, particularly in the clinical trials stage (Figure 15.1). The rationale is that identification of patients with increased risk of adverse drug reactions or drug inefficacy enables early tailoring of treatment so that a more optimal benefit–risk ratio is achieved at an individual level. This contrasts with overall population level drug information obtained typically from current randomized controlled trials (RCTs).

The use of biomarkers to define population subgroups for treatment is the basis of stratified medicine, the implementation of which requires early biomarker development, appropriate sample collection in clinical trials, and development of the most relevant assays to achieve clinically interpretable results [3]. In contrast to standard clinical trials, a pharmacogenomic or stratified medicine clinical trial seeks to determine if the treatment effect of a drug varies between subjects with different genotypes and the cost–benefit ratio of genotype-guided treatment over standard care. This chapter provides an introduction to clinical trial methodology and explores the specific considerations required for pharmacogenomics and stratified medicine.

A clinical trial is a planned experiment on humans designed to evaluate the effectiveness of one or more forms of treatment. Clinical trials can be broadly classified into five phases (Table 15.1), which represent stages in the clinical development of a drug. Phases 0–III are done before a license is granted and phase IV is done after authorization. Phases 0–III can last up to 10 years, and in that time a majority of drugs (~90%) are withdrawn from development due to safety issues, poor pharmacokinetic/pharmacodynamic (PK/PD) profiles, and poor efficacy in patients with the target disease [2].

15.2 PHASE I STUDIES

Following Phase 0, further exploratory studies after the initial first-in-human study consist of trials using multiple ascending doses of the drug (investigational medicinal product, or IMP) in a small group of subjects to assess safety, determine a safe dosage range, and assess interaction with

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00015-3
© 2014 Elsevier Inc. All rights reserved.

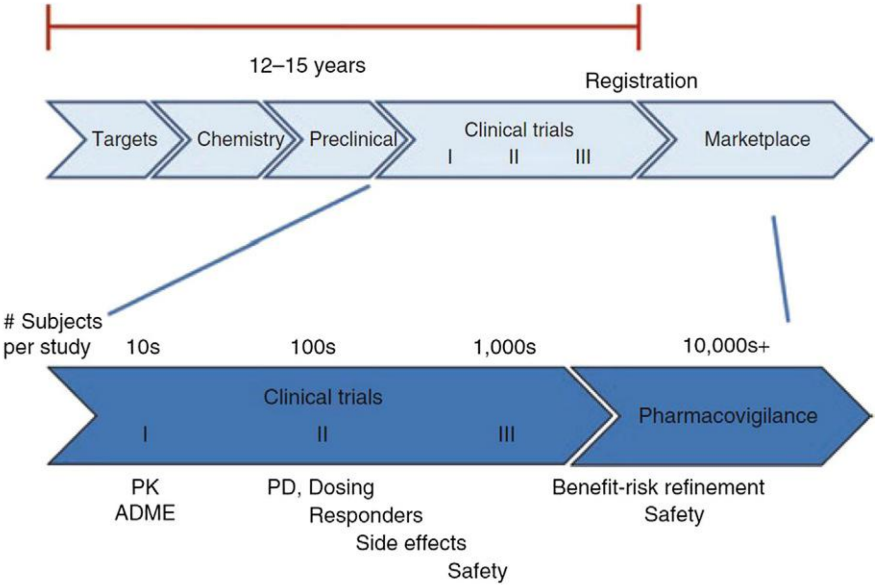


FIGURE 15.1 **Schematic representation of pharmaceutical research and development.** Examples of pharmacogenomic utility are shown beneath the stages of clinical development. ADME—adsorption, distribution, metabolism, and excretion; PD—pharmacodynamics; PK—pharmacokinetics. *Source: Reprinted with permission from Burns* [1].

TABLE 15.1 Clinical Trials

Phase	Objective
0	• First-in-human trials • Assessment of PK/PD, tolerability, and safety: single subtherapeutic doses of the study drug given to 10–15 subjects • Results with those of preclinical studies.
I	• Assessment of pharmacodynamics, pharmacokinetics, and toxicity. • Conducted in 50–200 healthy subjects or patients not expected to benefit from investigational product
II	• Initial clinical investigation of drug • Conducted in 100–400 subjects with target disease • Used to find best dose and investigate safety
III	• Full-scale evaluation of treatment • Generally major trials in (>1000–5000 subjects with target disease) • Aimed at conclusively demonstrating efficacy. • Typically studies on which registration of a new product based.
IV	• Post-marketing surveillance • Carried out after registration of a product • Can involve many thousands to millions of patients with target disease • Assess long-term safety of medicine (pharmacovigilance) and allow comparison with similar medicines. • Often for marketing purposes as well for broader experience with use of new product

food, gender, and other medicines. Furthermore, additional information on PK parameters of absorption, distribution, metabolism, elimination, and bioavailability are obtained using radiolabeled drug molecules; the effect on cardiac repolarization is assessed by reviewing changes in the QT interval in the electrocardiogram. Subjects who volunteer for phase I trials get no therapeutic benefit from the IMP, so appropriate and rigorous risk assessments are required, before the trial, to minimize the risk of harming the subjects. The overall safety of Phase I trials is good, with an incidence of 0.02% of serious adverse events related to the investigational drug [4]; however, there is paucity of safety data on nonchemical investigational drugs such as proteins, cytokines, and monoclonal antibodies.

There are several guidelines on calculating a safe starting dose on the IMP in Phase I trials [5,6]. The FDA [6] uses a stepwise model:

1. Convert the no-observed-adverse-effect dose level (NOAEL) from the toxicology studies to a human equivalent dose (HED) on the basis of body surface area.
2. Select a HED from the most appropriate species.
3. Apply a safety factor (10-fold) to give a maximum recommended starting dose (MRSD).
4. Adjust the MRSD on the basis of the predicted pharmacological action of the IMP.

An alternative or complementary method is to calculate the safe starting dose based on the minimal anticipated biological effect level (MABEL)[5] see also the Joint

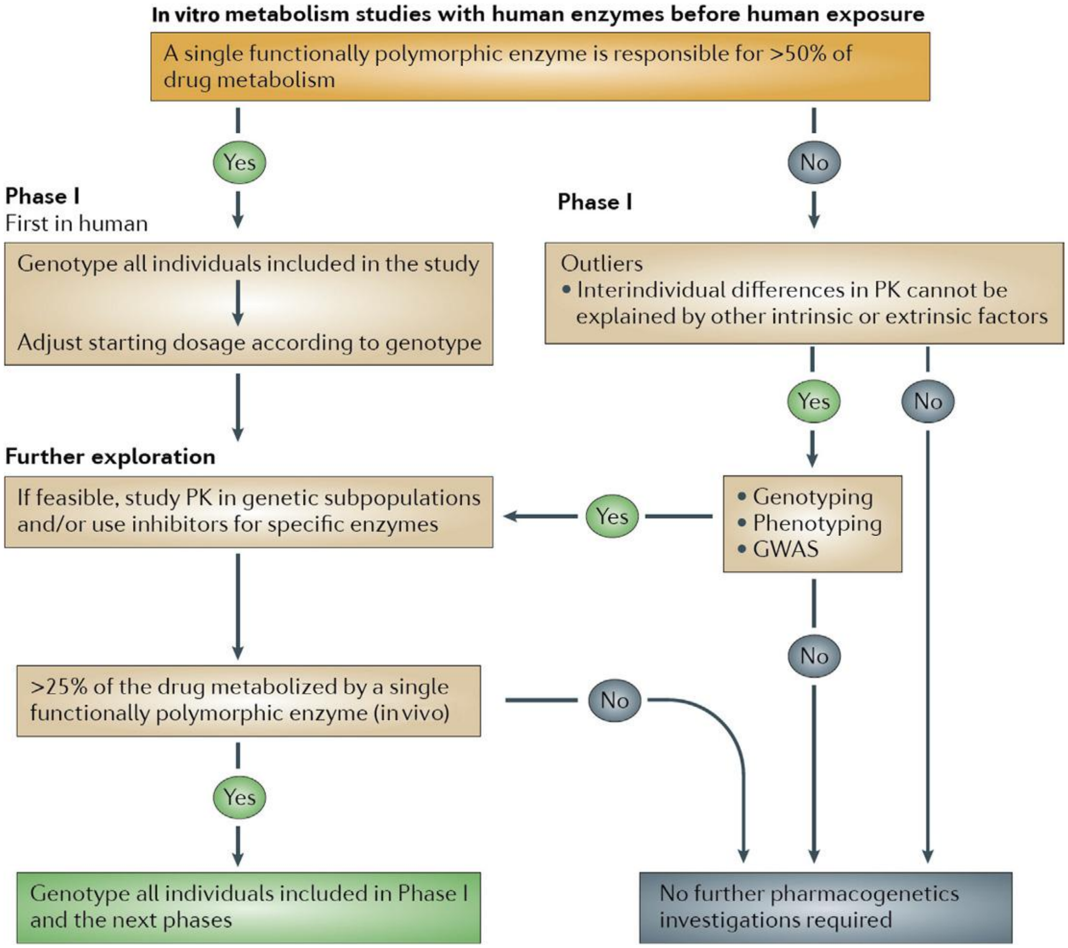

In vitro metabolism studies with human enzymes before human exposure
A single functionally polymorphic enzyme is responsible for >50% of drug metabolism
Yes
No
Phase I
First in human
Genotype all individuals included in the study
Adjust starting dosage according to genotype
Phase I
Outliers
• Interindividual differences in PK cannot be explained by other intrinsic or extrinsic factors
Yes
No
Further exploration
If feasible, study PK in genetic subpopulations and/or use inhibitors for specific enzymes
Yes
• Genotyping
• Phenotyping
• GWAS
No
>25% of the drug metabolized by a single functionally polymorphic enzyme (in vivo)
No
Yes
Genotype all individuals included in Phase I and the next phases
No further pharmacogenetics investigations required

FIGURE 15.2 EMA decision tree for *in vitro* studies prior to human exposure and Phase I studies. For polymorphic enzyme systems for which well-validated *in silico* physiologically based pharmacokinetic (PK) models have been developed, pharmacogenetic differences in humans may be predicted and used as a guide for clinical study design with respect to pharmacogenetics investigation. GWAS—genome-wide association study. *Source: Reprinted with permission from Maliepaard et al.* [11].

ABPI/BIA Taskforce Report, 2006. This approach uses all relevant information, including novelty, potency, mechanism of action, degree of species specificity, dose-response data from human and animal cells *in vitro*, dose- and concentration-response data from animals *in vivo*, pharmacokinetic and pharmacodynamic modeling, calculated target occupancy versus concentration, and target or target cell concentration in humans *in vivo*. In a dose-rising trial, the dose is usually doubled at each increment, a practice justified by a log-linear dose-response curve.

Studying healthy volunteers instead of patients may reduce variations caused by disease, advanced age, co-medications, and other factors. A limitation is that usually only surrogate parameters of efficacy or side effects can be measured. Nevertheless, these studies have played a major role in improving our understanding of pharmacogenetic polymorphisms such as those in CYP2D6, CYP2C19, thiopurine methyltransferase, and N-acetyltransferase type 2. Phase I studies in twins helped establish early heritability of drug pharmaockinetics [7].

Pharmacogenetic variations may be predicted from *in vitro* studies prior to the first application of the drug in humans. Both the FDA and the European Medicines Agency (EMA) have provided some guidance on integrating pharmacogenetics in early-phase clinical studies (Figure 15.2). The rationale for using pharmacogenetics in Phase I trials is primarily to facilitate successful drug development. For example, genetic variants can explain lack of efficacy or ADRs and prevent early failure of a new IMP. It may be much less expensive to include some pharmacogenetic analyses directly in phase I, rather than perform separate human pharmacogenetic studies later, because having the clinical pharmacogenetic knowledge before starting phase II allows for informed decisions about whether to stratify, exclude, or adjust according to pharmacogenetic diagnostics in phase II trials. Estimation of the relative impact of pharmacogenetic variation versus drug–drug interaction effects, effects of age, effects of body weight, and so forth, allows for the development of the correct priorities and exclusion criteria in phase II studies [8,9].

FDA guidance [10] emphasizes the importance of considering data from preclinical assessment in determining the best approach in Phase I studies. For example, after *in vitro* studies, clinical pharmacology studies may be used to identify any association of the IMP's PK/PD with genetic variants in metabolizing enzymes, drug transporters, and drug targets. If it is determined during preclinical testing that the IMP is metabolized by a well-established polymorphic gene (for example, cytochrome P450), the guidance recommends that the contribution of genetic variants to PK variability, and subsequently to dose or dosing regimen selection in human studies, should be evaluated.

The EMA has also developed detailed guidelines on assessing pharmacokinetic pharmacogenetics in Phase I studies of novel drugs [11]. When the *in vitro* data indicate that >50% of the drug is cleared by a single functionally polymorphic enzyme, it is recommended that relevant gene be genotyped in the first-in-human (Phase 0) study population to avoid safety issues related to genetically determined differences in active substance exposure. Subjects with a poor metabolizer genotype (which can result in higher levels of the active drug or its metabolites) are allowed to participate in the first-in-human study only at doses lower than those expected to be safe in individuals who have an extensive metabolizer phenotype. Further exploration studies require investigation of the relative contribution of an important polymorphic enzyme on the *in vivo* PK properties of a drug or active metabolite, usually in a conventional PK study within genetically defined subpopulations whenever feasible [11]. When this is impractical, and when there is ample supporting scientific literature or validated data available, the effect of a genotype may be confidently mirrored by treatment with an inhibitor of the respective metabolizing enzyme; the effect of the polymorphism can then be estimated from the results. In such cases, it is important to consider off-target effects of the respective inhibitors. If a marked effect of a polymorphism is confirmed *in vivo* (arbitrarily defined in the EMA guideline as when >25% of the parent drug is cleared by the polymorphic enzyme *in vivo*), genotyping the indicated genes is recommended in as many of the Phase I studies and the next phases as possible. This is to maximize the amount of supporting data for recommendations utilized in studies of genetically defined subpopulations [11,12].

The FDA considers any level drug metabolism by a single polymorphic enzyme to be important, because it is relevant to drugs with steep exposure–response relationships. (This is in contrast to the EMA, which uses a threshold of >50% of the drug to be cleared by the enzyme in *in vitro* studies). When *in vitro* studies suggest that a drug is metabolized extensively by a polymorphic enzyme, FDA guidance recommends that pharmacogenetic analyses be conducted in single- and/or multiple-dose PK studies to evaluate the impact of common genetic variants on *in vivo* drug concentrations (and relevant PD measurements if available). Consideration of known differences in the prevalence of ADME-related gene variants among racial or ethnically distinct groups should be incorporated into the planning for such analyses. Selected genetic variants (for example, SNPs in genes encoding metabolizing enzymes) that occur with acceptable prevalence in the population should be studied. A dedicated clinical pharmacology study with targeted, genotype-based enrollment may be desirable in certain situations and should allow a meaningful retrospective analysis [12].

If a drug is converted into an active metabolite through a polymorphic metabolic pathway, rather than a candidate gene study, other strategies that include the screening of subjects in early-phase clinical trials using high-throughput methodologies (for example, ADME "gene chips") may generate valuable information on the sources of PK and/or PD variability.

15.3 PHASE II STUDIES

Phase II studies are dose-finding investigations performed to aid the design of subsequent Phase III studies, including the decision of whether or not genotype-based dosing should be applied. **The standard phase II design is** a single-arm study that tests the null hypothesis of insufficient efficacy versus sufficient activity meriting further investigation. A type I error defines the chance that an ineffective agent will be studied further; a type II, error the chance that an effective agent will not be studied further, which is usually regarded as the more serious error in phase II testing. Phase II clinical trials are able to detect only a large treatment improvement (e.g. greater than 10%) because the sample size is small. An important characteristic of some phase II trial designs is the use of early stopping rules. If there is sufficient evidence that one of the treatments under study has a positive treatment effect, patient accrual is terminated and the treatment is declared promising. Also, if a treatment is sufficiently shown not to have a desirable effect, patient accrual is terminated and the treatment is declared not promising.

A **randomized phase II selection design** allows multiple single-arm studies to be conducted in the same timeframe and with the same entry criteria [13]. Typically, this randomizes between two or more experimental arms without a control arm [3,4,8]. Usually, there will be a test for the activity of each arm, using standard criteria for single-arm studies, and a rule for selecting the "best" of the active arms for further testing. Often, the rule for a selection design study is to select as "best" the arm with the best efficacy level, no matter what the magnitude of the difference. There is generally not adequate power for formal tests comparing arms. A weakness of this design is the reduced likelihood of being able to select the best arm if there are an increasing number of arms or if there is a small difference in activity among them [14]. The advantages of a randomized study over separate studies include decreasing the effects of patient selection bias, population drift, and stage migration, and the ability to ensure that uniform evaluation criteria are used.

Whereas a randomized selection design may have advantages when screening multiple treatments, a **randomized controlled phase II study** may be useful when fewer comparisons are required. This design typically compares an experimental arm to a control arm (with or without a placebo). In settings where an effective therapy exists, the control arm might consist of effective standard therapy plus placebo. The added expense of a placebo and the administrative burden need to be weighed against the potential for bias. The effectiveness of the placebo in blinding treatment should also be considered because many agents have significant side effects.

Comparison to a control arm is most useful when there is little prior information on expected efficacy rates in a population; it can also be useful for end points, which can be heavily influenced by patient selection, such as time to progression end point (TTP) and progression-free survival in cancer studies [15]; and when it is expected that improvements in supportive care or other clinical factors will change the expected outcome in the population being studied. The efficacy end point can be a standard measure of disease status such as objective response or TTP, allowing the study to be completed with fewer patients than required in a phase III study of survival. Because the primary end point is not survival, a crossover to the drug from the placebo arm when the end point is reached can also be considered; it is advantageous because it attracts patient accrual. However, its weakness is that it can dilute determination of a survival benefit.

For evaluating TTP in a study with crossover at progression, there is usually a serious risk of bias, especially when the control arm includes no active agents. However, in many phase II studies the magnitude of the potential bias should be small relative to the effect that can be detected. Type I and II error rates vary, but as in other phase II designs, the type I error can generally be larger than in phase III—Korn et al. [15] suggest considering one-sided type I error rates as large as 20%. The magnitude of the difference between the null and alternative hypotheses may also be larger than appropriate for a phase III study. A potential weakness of a randomized phase II control design is the need for a second larger study. A decision as to whether a single-arm phase II study will be adequate to make a decision to go forward to a phase III study, or whether a randomized comparison with a progression end point will be required, should be made prospectively.

A **randomized discontinuation design** is an example of an enrichment strategy [16] in which patients are randomized between continuing the drug or going on a placebo if free of progression at some defined time point. Patients are usually crossed over from placebo to treatment at progression or at a specified progression-free interval. Those who are responding continue on therapy until progression. The design is complex, with three registration points: initial registration, at the defined time of randomization during stable disease, and crossover at progression. It is most attractive in settings where it is thought that only a subset of the population will benefit from the treatment, which might be the case for some targeted therapies. It has been shown that while theoretically effective in these settings, their effectiveness depends on the extent to which the initial run-in can select out the subgroup that is clearly benefiting on the basis of disease stabilization versus rapid progression [17]. Potential weaknesses of this study design include the possibility of a carry-over effect from the run-in that could dilute differences between the randomized arms. Desired patient numbers also need to be considered carefully because these studies tend to be much larger than other phase II studies.

The EMA recommends that potential clinically relevant pharmacogenetic influences on drug PK properties from Phase I studies be considered in the design of Phase II studies, either by investigating genotype-based or phenotype-based dosing or by exploring therapeutic drug monitoring (TDM) or biomarker-guided dosing (Figure 15.3) [11,12]. When Phase II data indicate that the difference in exposure observed between genetically defined subpopulations is likely to be of clinical importance, individuals with an intermediate metabolizer phenotype should be investigated in a further PK study. If the sponsor decides that dose adjustment is not needed based on genotype, TDM, or biomarker data, the exposure level obtained with an unadjusted dose in the genotypically defined subpopulation should also be used in Phase II studies [11,12].

FDA guidance recognizes that important PK variability of an active parent drug (or metabolites), if observed in Phase I trials, should shape the design of subsequent clinical studies (for example, dose–response studies in genetically defined subpopulations). For cases in which genotypes are found to be relevant in predicting exposure and drug effect, this knowledge should be used in the design of subsequent clinical trials. For example, genotypes can be used to select patients for trials (e.g., enrichment with responders or exclusion of patients likely to experience toxicity); stratify groups within trials; and adjust doses for trials. This strategy may increase average effect, decrease toxicity, and improve the chances for overall study success.

Another important aspect of the design of clinical pharmacology studies is the pharmacogenetic evaluation of

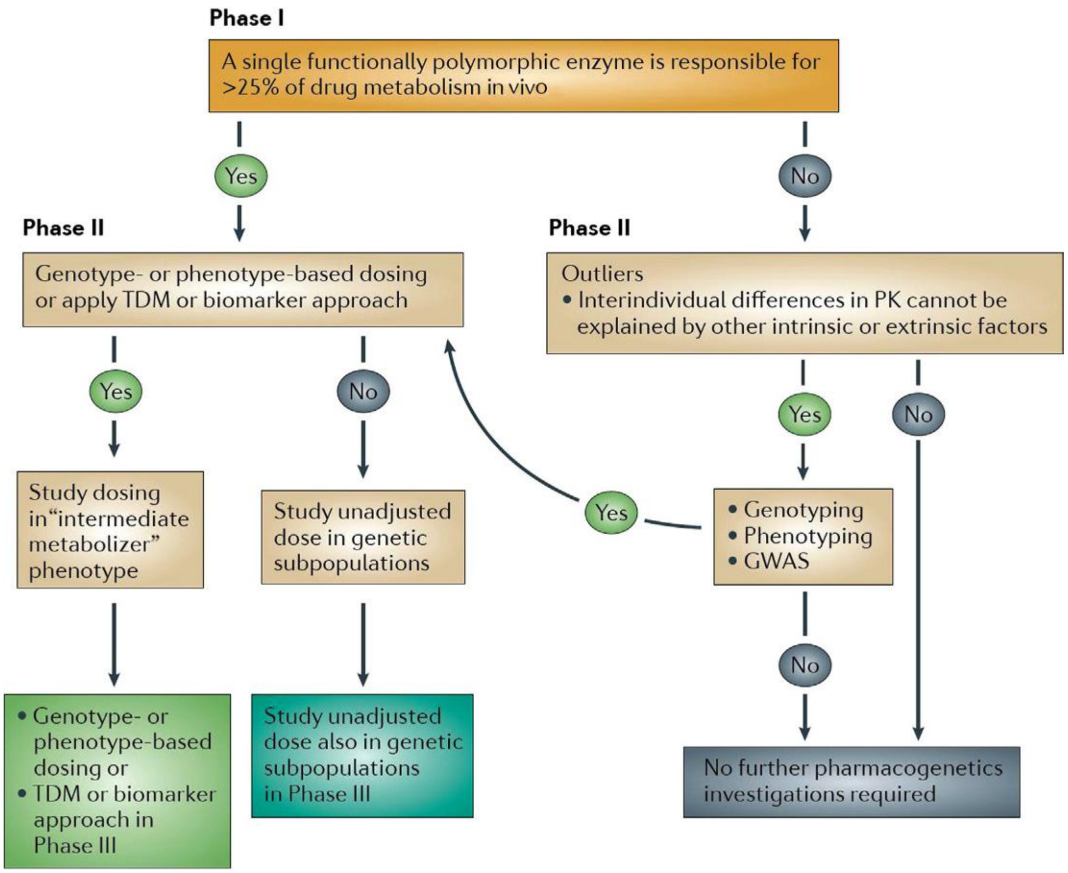

Phase I
A single functionally polymorphic enzyme is responsible for >25% of drug metabolism in vivo
Yes
No
Phase II
Genotype- or phenotype-based dosing or apply TDM or biomarker approach
Phase II
Outliers
• Interindividual differences in PK cannot be explained by other intrinsic or extrinsic factors
Yes
No
Yes
No
Study dosing in "intermediate metabolizer" phenotype
Study unadjusted dose in genetic subpopulations
Yes
• Genotyping
• Phenotyping
• GWAS
No
• Genotype- or phenotype-based dosing or
• TDM or biomarker approach in Phase III
Study unadjusted dose also in genetic subpopulations in Phase III
No further pharmacogenetics investigations required

FIGURE 15.3 **EMA decision tree for Phase I and Phase II studies.** *Source: Reprinted with permission from Maliepaard et al.* [11].

dose–response relationships. If the genotype is shown to be associated with systemic exposure–response (activity and/or safety) in early PK or PD studies, subsequent dose–response studies with methodological features including, but not limited to, genotype-based dose stratification should be considered.

Ideally, pharmacogenetic variants relevant to drug metabolism and transport are already included in late Phase II trials to enable PGDx-based dose adjustments in subsequent Phase III trials. However, if the population frequency of a typical pharmacogenetic polymorphism is >50%, then it is clear that pharmacogenetic questions can only be solved in larger Phase III trials.

15.4 PHASE III STUDIES

Phase III trials are also called comparator trials or controlled trials, as they involve direct comparison between two or more treatments (Figure 15.4). In this phase, data are gathered from large numbers of patients (hundreds to many thousands) to find out whether the new treatment is more effective and possibly has fewer side effects than the current standard treatment. These clinical trials take many years to complete. Although phase III clinical trials focus on patients with a specific disease, they typically include patients of different ages, ethnicities, and genders so that the results will be applicable to a large number of people.

A major problem is the possibility of bias distorting the results of the study, and studies are rigorously planned to minimize it. Usually a treatment or a sequence of treatments is assigned to each group and subjects are followed up prospectively and simultaneously. The former case is referred to as a parallel group design; the latter, as a cross-over design. Three types of bias commonly observed:

- Selection bias is defined as a nonrandom imbalance among treatment groups of the distribution of factors capable of influencing the end points—that is, of subexperimental factors (including prognostic factors). This is mitigated using randomization (see Box 15.1).
- Assessment bias is defined as a nonrandom imbalance among treatment groups in the way subjects are followed and assessed during the study. It can be mitigated using suitable blinding procedures (see Box 15.2).

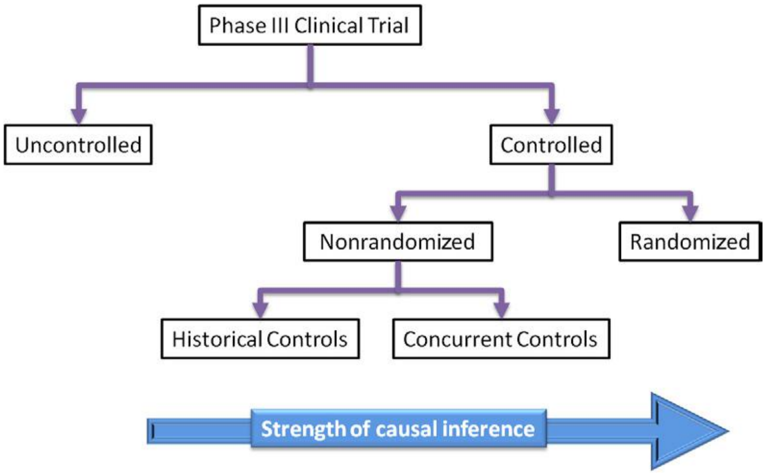


FIGURE 15.4 Phase III clinical trials.

BOX 15.1 Randomization

Randomization is a process that assigns research patients by chance, rather than by choice, to either the treatment group or the control group. This implies that each individual assignment cannot be predicted based on previous assignments. Put simply, at each assignment each subject has the same probability of being assigned to each treatment: this probability is 50% in a study with two treatments, 33.3% in a study with three treatments, and 25% in a study with four treatments. In addition to minimizing selection bias, by introducing a random element into the assignment of selected subjects, randomization allows the application of the statistical significance test.

In *simple randomization*, each patient has the same probability of receiving each of the study treatments (or sequence of treatments). This is usually achieved by tossing a coin or generating random numbers.

Permuted block randomization is a form of restricted randomization in which blocks of k patients are created such that balance is enforced within each block. One of the blocks is then selected at random and the k patients are assigned accordingly. This promotes group balance at the end of the trial, as well as periodic balance in the sense that sequential patients are distributed equally between groups. This type of randomization can be susceptible to selection bias.

In *stratified randomization*, if a covariate (e.g., age, gender, race, or center) is known to be the cause of heterogeneity among patients, patients are stratified into several homogeneous strata with respect to it. Randomization is then performed within each stratum (usually blocked). Stratified randomization guarantees treatment balance within risk factors. An extreme case of stratification is matching (common in case-control studies).

BOX 15.2 Blinding

Blinding/masking is used to prevent research outcomes from being influenced by either the placebo effect or observer bias. It does so by blocking the identity of the treatments. Blinding/masking embraces a multiplicity of situations in which the researcher systematically favors patients belonging to one treatment group over another when carrying out procedures linked (directly or indirectly) to the assessment of response to treatment. Commonly, a study is defined as double-blind when neither the research staff nor the patients know which study treatment is administered to an individual patient. When only the patients are unaware of the study treatment, the study is defined as single-blind. A study in which there is no blinding is defined as open-label. Triple-blind studies are those in which none of the patients, investigators, or sponsors know the treatment.

Phase III Study Designs

- Randomized Control Trials (parallel design)
- Uncontrolled Trials (single-treatment)
- Historical Controls
- Nonrandomized
- Concurrent Trials
- Crossover Designs
- Factorial Designs
- Group Sequential Design

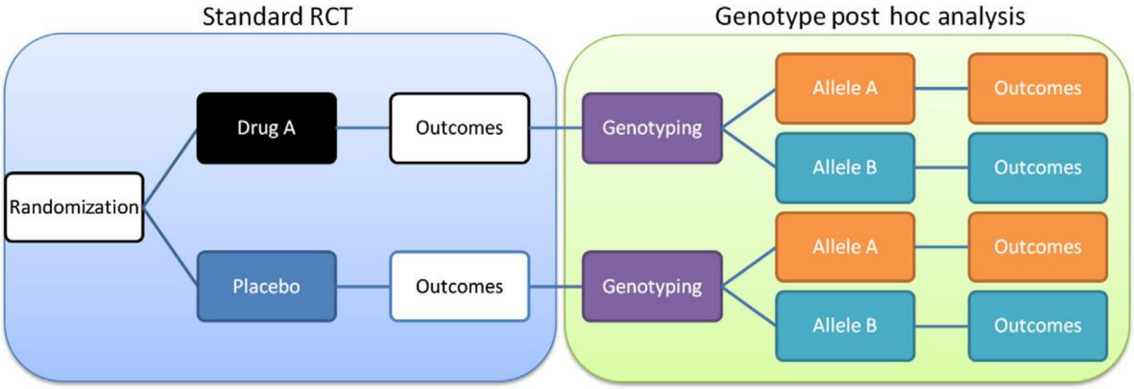


FIGURE 15.5 **Randomized trial with post hoc analysis.**

- Analysis bias is defined as a distortion in favor of one of the treatments, intervening during the data analysis, and is minimized using an *a priori* definition of the statistical methods and populations.

15.4.1 Randomized Controlled Trials

Randomized controlled trials (RCTs) are the gold standard for testing new medicines; they have an important role in determining the efficacy and safety of interventions. An RCT is characterized by shared principles: randomization of subjects into treatment and control groups; blinding of all participants as to which group subjects are in; the need for follow-up; and the handling of dropouts and withdrawals (see Boxes 15.1 and 15.2). Within an RCT, a subject will be randomly allocated to a group that receives one of the treatment options or a placebo. Typically, each group contains the same distribution of relevant demographics, such as age, sex, and ethnicity. RCTs are typically time-consuming, and large sample sizes are often needed to ensure sufficient statistical power. Still, they at best demonstrate the small relative benefit of a 15–20% relative reduction in a composite clinical end point.

RCTs may be classified into two subtypes depending on what they seek to prove [18,19]. Explanatory RCTs (ERCTs) evaluate the efficacy of a pharmaceutical agent under highly controlled conditions that do not necessarily reflect real life. Pragmatic RCTs (PRCTs) assess the effectiveness of a treatment in everyday situations, meaning that a less than strict adherence to the treatment regimen is not a reason for subject exclusion. ERCTs can produce more precise assessments of a pharmaceutical agent, although the trial design often deviates significantly from daily routine and the results are suited to the “average patient,” who does not exist. Also, they are not ideal for tailoring treatment to an individual, and without prior consideration of the underlying genetics, patients in a particular trial group might receive a drug that is ineffective for their genotype or, worse, dangerous [18]. PRCTs, on the other hand, generate weaker evidence but better reflect the circumstances of patients as they go about their daily lives.

Pharmacogenomic clinical trials can be designed (or analyzed post hoc) with the intention to study whether a subgroup of patients, defined by certain genetic characteristics, responds differently to the treatment (Figure 15.5). Alternatively, a trial can be intended to verify whether genotype-guided treatment is beneficial over standard care. For example, gefitinib, a lung cancer drug, inhibits epidermal growth factor receptor (EGFR) signaling in target cells, providing patients with a significantly longer period without disease progression and fewer side effects than with conventional chemotherapy [20]. Only patients with a specific mutation in the gene that encodes EGFR receive any benefit from gefitinib; however, initial clinical trial results were negative because none of the subjects had been selected based on EGFR activity. Positive clinical results were seen only after these subgroups were identified [21,22].

Using data from a completed trial in which DNA was already gathered at baseline and stored (as is more and more the standard practice in Phase III trials), a PGx subgroup analysis is relatively simple, fast, and inexpensive. If the outcome event is independent of the gene under study, which means that the gene is not a prognostic factor for the outcome, it is valid for the trial to be what is known as an exposure-only. Then only data from treated patients are taken into account, which is justified because the risk for the outcome is identical for all untreated patients regardless of their genotype. The Exposure-only design has been applied in the study of cardiovascular outcomes in subjects with acute coronary syndromes treated with clopidogrel, with the assumption that there is no direct association between CYP alleles and cardiovascular outcomes [23].

The major disadvantages of a subgroup analysis include the following: type I error may increase when analyzing many subgroups; confounding bias can be introduced if the subgrouping variable is associated with another variable that also modifies the treatment effect; and genetic differences in allele frequencies between subgroups (population stratification) may lead to misleading conclusions if the distribution of these populations is unequal between the strata. In the case of pharmacogenetic subgroups, the chance of introducing bias is small since inheritance of a specific genotype is in general independent of inheritance of other traits. However, bias may be caused by selective consenting for genotyping by participants. At the start of the study, this phenomenon is not likely to introduce bias because consenting is not related to study outcome. If, however, DNA or other biomaterial is collected (which can be reused for post hoc genotyping), the availability of the material might be selective owing to more frequent use from exceptional patients. If there is no DNA collected at baseline and it needs to be collected post hoc, a bias might be introduced by a selective loss of patients or by selective consenting by participants, because the study's outcomes can influence the decision to consent.

With an **enrichment design**, all potentially eligible patients are first tested for the genotype or biomarker and only biomarker positives are randomly assigned to the experimental or control treatment (Figure 15.6). Other patients are in principle excluded from further investigation in the study. This is similar to the post-hoc study design, except the subgroup analysis is planned a priori. If the prevalence of the selected genetic variants is unequal, it is possible to selectively include participants to create balanced numbers in the genotype strata [24].

The pivotal trial for trastuzumab is a well-known example of an enrichment design [25] in which patients with HER2-positive breast adenocarcinoma (human epidermal growth factor receptor 2, or HER2) were enrolled and randomly allocated to chemotherapy with or without trastuzumab. This study provided strong evidence that trastuzumab combined with chemotherapy improves outcomes among women with HER2-positive breast cancer.

With genotype information available before randomization, it is also possible to select patients with genotypes between which the largest difference in treatment effect is expected, instead of including all genotypes. In a "classical" enrichment design, only patients expected to show the greatest benefit from the intervention are enrolled. Using pharmacogenomics information in the selection is also possible by excluding genotypes assumed to be associated with an increased risk for side effects or nonresponse. This strategy reduces sample size and can increase subject safety; thus, it may eliminate the need for other safety measures—for example, close monitoring of drug plasma concentrations during the trial [26]. Its main advantage is that sample size is determined taking the planned subgroup analysis into account, which decreases the risk of insufficient statistical power at the end.

Bias introduced by selective consenting is avoided by genotyping at the beginning of the study as well. Moreover, the possibility of randomization by genotype, creating balanced numbers of genotypes per treatment arm, further increases the chances of a significant result and minimizes the possibility of confounding introduced by unequal genotype distribution between subgroups [27]. Selective inclusion of particular genotypes to create similar numbers of subjects in the genotype strata may prolong recruitment time if a prespecified number of patients carrying a rare allele are to be included. If time is a limiting factor, it might be a better option to settle for fewer patients with the rare genetic variant—a genotype ratio of 2:1 instead of 1:1, for example—which results in only minimal loss of power. Helpful in this consideration is the number needed to genotype in order to find an eligible patient; this number is calculated by dividing one by the prevalence of the genotype(s) under study [28].

When the existing evidence is not strong enough to exclude a subgroup of patients based on genotype, an **adaptive trial design** provides an alternative. In this design, the accumulation of data is used to modify aspects of the study because it continues without undermining the trial's validity and integrity.

Adaptive clinical trial designs are likely to be a vital component of drug development for multiple-target therapeutics [29]. Such designs, exemplified by I-SPY2 ("investigation of serial studies to predict your therapeutic response with imaging and molecular analysis 2") and BATTLE ("biomarker-integrated approaches of targeted therapy for lung cancer elimination"), allow researchers to review data during the clinical trial and adapt patient recruitment accordingly. Depending on early response assessment, such trials might allow researchers to test patient cohorts selected for multiple potential biomarkers for different therapeutic approaches and prioritize patients for distinct arms of the trial.

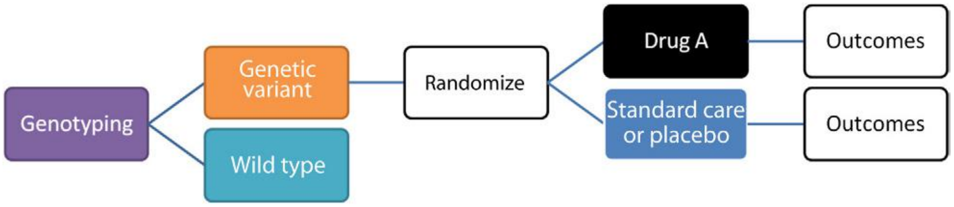


FIGURE 15.6 **Enrichment design.**

Designs such as the adaptive clinical trial require robust and advanced statistical techniques and may heighten the risk of false-positive conclusions. Particular care must be given to determining whether a putative predictive biomarker may also have prognostic relevance, where patient cohorts may have a superior outcome determined by the biomarker used to select them rather than the systemic therapy itself. Because of the requirement for interim analyses and clinician unblinding, traditional double-blind standards may be affected [29].

It was only in 2007 that the EMA published a paper outlining the appropriate conduct for adaptive-trials, and only in February 2010 that the FDA followed suit with its draft guidance to industry [29]. By changing the patient population to one that is more likely to benefit from a treatment, adaptive trials can inflate the risk of reaching a false-positive conclusion. They also run the risk that the interim data required to take that step to modify the trial might compromise it's double-blind safeguards, influence patient and investigator behavior, and color the results even further. For these reasons, the FDA and the EMA advise against using an adaptive trial when a standard trial will do, and urge extra caution in designing the late-stage clinical trials that are crucial to determining drug approvals.

In the **parallel controlled pharmacogenetic diagnostic study** (Figure 15.7), randomization is performed either in genotype-guided treatment or in standard care [8,27]. Genotype-guided treatment can include an alternative treatment (or exclusion) for subjects with predetermined genotypes. Or an algorithm that includes a subjects' genotype (and possibly other parameters such as age and weight) can be used to determine the initial or maintenance dose in this treatment arm. Analysis of this RCT is similar to that in a "regular" RCT, in which two interventions are compared. The advantage of parallel controlled studies is that their results deal with some of the main barriers to clinical uptake: skepticism toward clinical evidence and a perception that alternative ways of handling the risks of ADRs suffice [8,27]. Also, methodological or logistical challenges in implementing genotyping in daily practice may arise during the study that were not considered outside the trial setting.

One of the first RCTs to use this study design, randomly assigned HIV-infected patients to either prospective HLA-B*57:01 screening or standard-care use of abacavir [30]. The results, combined with results from previous observational studies, were reason enough to change abacavir's FDA drug label. Randomization to either standard care or pharmacogenomics-based treatment, enables a direct estimate of the benefit that genotyping provides to the prevention of ADRs or nonresponse. This estimate is necessary to assess the cost-savings of genotyping [31]. At the same time, a "break even point" of the benefit can be calculated, illustrating the number of patients needed to genotype, and treat accordingly, for the testing to be (cost-) effective.

In the "randomize-all" or "all-comers" study design, all patients meeting trial eligibility criteria, irrespective of their biomarker or genotype status, are randomly allocated to either experimental or control treatment, after which associations between biomarker status and treatment response are evaluated. This study design has the same patient flow elements as RCTs for evaluating treatment options. Alternatively, a biomarker can be evaluated retrospectively, using data and stored biospecimens collected from previously completed RCTs. In such scenarios, trials are commonly labeled "biomarker analysis within an existing RCT" or "prospective/retrospective" [32,33]. Type of randomization is another source of variability in labeling of randomize-all trials. In cases where a simple 1:1 randomization procedure is applied to all patients, trials are labeled "simple randomization." However, in cases where the biomarker under evaluation is binary or categorical with few categories, randomization can be done separately for each biomarker category through stratified randomization. Labels such as "biomarker-stratified," "stratified randomized," "nontargeted RCT (stratified by marker)," "stratified analysis," "stratification," and "separate randomization," all refer to this randomization type [34]. A special case of stratified randomization is randomization performed by means of a Bayesian response-adaptive procedure, rather than a standard equal-randomization procedure; this is known as "Bayesian adaptive randomization design" or "outcome based adaptive randomization"[35].

Another study design is the biomarker strategy in which biomarker status is assessed in all patients enrolled in the trial, who are then randomly allocated to either the biomarker strategy arm or to the standard treatment arm. There

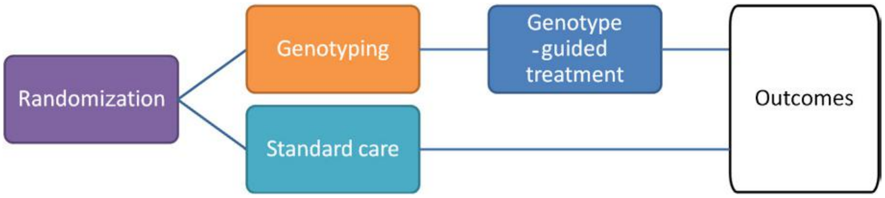


FIGURE 15.7 **Parallel controlled pharmacogenetic diagnostic study**.

is also modified marker-based strategy design, in which a second randomization between experimental and control therapy replaces the control arm [34].

A modification of the pharmacogenomic cohort study frequently preferred is the **nested case-control study**, also known as a **case-cohort study**. It is not necessary to analyze all subjects of the RCT; typically, selection is of all available participants of the study who had the outcome of interest (cases) and randomly selected participants who did not have the outcome of interest (controls). Typically, one control per case is chosen, but power can be increased by including up to four controls per case. Although the term "nested case-control" refers to a retrospective design, this type of study is indeed prospective in nature. It differs from retrospective case-control studies because it is only a random selection from a prospective cohort and therefore many serious types of bias known to arise in case-control studies are absent. A recent example of this study type involved statin-induced myopathy [36].

One of the drawbacks of RCTs is that they are conducted in artificial environments where significant attempts are made to control all factors that can influence outcome. Thus, an artificial increase in the standard of care may result in dilution of the effect of a biomarker in decision making. In addition, it has been suggested that the simple fact of participating in a RCT may have a positive effect on outcome, which may further complicate the generalization of the findings from a RCT to the general population [37].

15.5 PHASE IV TRIALS

An important source of data for pharmacovigilance is individual case reports (or case series) related to ADRs. Fatal outcomes reported in carriers of certain inherited variants may stimulate further research and attract increased funding for pharmacogenomic studies. Case reports do not provide generalizable results, which is a major limitation.

15.6 RETROSPECTIVE ANALYSIS OF CLINICAL DATASETS

Retrospective subgroup analyses have proven valuable in defining distinct patient subgroups that benefit preferentially from targeted therapeutic approaches. Examples include EGFR monoclonal antibody therapy benefit and KRAS wild-type status [38], HER2 status and paclitaxel benefit [39], and EGFR mutation status and gefitinib survival benefit [40]. Retrospective analyses have proven useful to several functional genomics RNA interference strategies in defining the relevance of a putative biomarker to treatment response in defined breast cancer subtypes [41–43].

Retrospective molecular analyses of clinical trial tissue datasets can prove fruitful in determining the predictive relevance of previously defined molecular aberrations, such as the importance of chromosome 17 centromeric enumeration probe duplication and anthracycline benefit in the adjuvant therapy of primary breast cancer [44]. The case–control design is the most frequently applied observational retrospective study in PGx and in genomic disease susceptibility research. Persons identified as having specific ADRs or poor therapy outcomes are the "cases." The frequencies of genotypes in "cases" are compared with those in concurrently sampled controls who have had comparable drug exposure but are free of problematic outcome. It is likely that, in the future, an acceleration in pharmacogenomic applications will be achieved by combining information from RCTs and observational studies in which there is no concurrent random assignment comparing a population without intervention to a population with intervention. This is a rapidly evolving field, and a sound understanding of different study designs is crucial in transforming pharmacogenomic discoveries into successful clinical applications.

REFERENCES

[1] Burns DK. Developing pharmacogenetic evidence throughout clinical development. Clin Pharmacol Ther 2010;88(6):867–70.

[2] Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 2004;3(8):711–5.

[3] Bienfait KL, Shaw PM, Murthy G, Warner AW. Mobilizing pharmacogenomic analyses during clinical trials in drug development. Pharmacogenomics 2013;14(10):1227–35.

[4] Sibille M, Donazzolo Y, Lecoz F, Krupka E. After the London tragedy, is it still possible to consider Phase I is safe? Br J Clin Pharmacol 2006;62(4):502–3.

[5] Muller PY, Milton M, Lloyd P, Sims J, Brennan FR. The minimum anticipated biological effect level (MABEL) for selection of first human dose in clinical trials with monoclonal antibodies. Curr Opin Biotechnol 2009;20(6):722–9. <www.abpi.org.uk/our-work/library/guidelines/Pages/phase-1-trials.aspx>

[6] FDA. Estimating the safe starting dose for healthy volunteer. <www.fda.gov/downloads/Drugs/Guidances/UCM078932.pdf>. 2005.

[7] Vesell ES, Page JG. Genetic control of dicumarol levels in man. J Clin Invest 1968;47(12):2657–63.

[8] Stingl Kirchheiner JC, Brockmoller J. Why, when, and how should pharmacogenetics be applied in clinical studies? Current and future approaches to study designs. Clin Pharmacol Ther 2011;89(2):198–209.

[9] Kalow W, Tang BK, Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. Pharmacogenetics 1998;8(4):283–9.

[10] FDA. Guidance on clinical pharmacogenomics: premarketing evaluation in early phase clinical studies. <www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM243702.pdf>. 2011.

[11] European Medicine Agency. Guideline on the use of pharmacogenetic methodologies in the pharmacokinetic evaluation of medicinal products. <www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/02/WC500121954.pdf>. 2009.

[12] Maliepaard M, Nofziger C, Papaluca M, Zineh I, Uyama Y, Prasad K, et al. Pharmacogenetics in the evaluation of new drugs: a multiregional regulatory perspective. Nat Rev Drug Discov 2013;12(2):103–15.

[13] Gray R, Manola J, Saxman S, Wright J, Dutcher J, Atkins M, et al. Phase II clinical trial design: methods in translational research from the Genitourinary Committee at the Eastern Cooperative Oncology Group. Clin Cancer Res 2006;12(7 Pt 1):1966–9.

[14] Simon R, Wittes RE, Ellenberg SS. Randomized phase II clinical trials. Cancer Treat Rep 1985;69(12):1375–81.

[15] Korn EL, Arbuck SG, Pluda JM, Simon R, Kaplan RS, Christian MC. Clinical trial designs for cytostatic agents: are new approaches needed? J Clin Oncol 2001;19(1):265–72.

[16] Rosner GL, Stadler W, Ratain MJ. Randomized discontinuation design: application to cytostatic antineoplastic agents. J Clin Oncol 2002;20(22):4478–84.

[17] Freidlin B, Simon R. Evaluation of randomized discontinuation design. J Clin Oncol 2005;23(22):5094–8.

[18] Swanton C, Caldas C. From genomic landscapes to personalized cancer management-is there a roadmap? Ann NY Acad Sci 2010;1210:34–44.

[19] Liu L, Leung EL-H, Tian X. Perspective: the clinical trial barriers. Nature 2011;480(7378):S100.

[20] Douillard JY, Shepherd FA, Hirsh V, Mok T, Socinski MA, Gervais R, et al. Molecular predictors of outcome with gefitinib and docetaxel in previously treated non-small-cell lung cancer: data from the randomized phase III INTEREST trial. J Clin Oncol 2010;28(5):744–52.

[21] Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350(21):2129–39.

[22] Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304(5676):1497–500.

[23] Mega JL, Close SL, Wiviott SD, Shen L, Hockett RD, Brandt JT, et al. Cytochrome p-450 polymorphisms and response to clopidogrel. N Engl J Med 2009;360(4):354–62.

[24] Kemmeren JM, Algra A, Meijers JC, Tans G, Bouma BN, Curvers J, et al. Effect of second- and third-generation oral contraceptives on the protein C system in the absence or presence of the factor VLeiden mutation: a randomized trial. Blood 2004;103(3):927–33.

[25] Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001;344(11):783–92.

[26] Murphy MP, Beaman ME, Clark LS, Cayouette M, Benson L, Morris DM, et al. Prospective CYP2D6 genotyping as an exclusion criterion for enrollment of a phase III clinical trial. Pharmacogenetics 2000;10(7):583–90.

[27] van der Baan FH, Klungel OH, Egberts AC, Leufkens HG, Grobbee DE, Roes KC, et al. Pharmacogenetics in randomized controlled trials: considerations for trial design. Pharmacogenomics 2011;12(10):1485–92.

[28] Mulder H, Heerdink ER, van Iersel EE, Wilmink FW, Egberts AC. Prevalence of patients using drugs metabolized by cytochrome P450 2D6 in different populations: a cross-sectional study. Ann Pharmacother 2007;41(3):408–13.

[29] Editorial. Time to adapt. Nature 2010;464(7293):1245–46.

[30] Mallal S, Phillips E, Carosi G, Molina JM, Workman C, Tomazic J, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008;358(6):568–79.

[31] Rodriguez-Antona C, Gurwitz D, de LJ, Llerena A, Kirchheiner J, de Mesa EG, et al. CYP2D6 genotyping for psychiatric patients treated with risperidone: considerations for cost-effectiveness studies. Pharmacogenomics 2009;10(4):685–99.

[32] Mandrekar SJ, Sargent DJ. Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges. J Clin Oncol 2009;27(24):4027–34.

[33] Wang SJ, O'Neill RT, Hung HJ. Statistical considerations in evaluating pharmacogenomics-based clinical effect for confirmatory trials. Clin Trials 2010;7(5):525–36.

[34] Tajik P, Zwinderman AH, Mol BW, Bossuyt PM. Trial designs for personalizing cancer care: a systematic review and classification. Clin Cancer Res 2013;19(17):4578–88.

[35] Zhou X, Liu S, Kim ES, Herbst RS, Lee JJ. Bayesian adaptive design for targeted therapy development in lung cancer–a step toward personalized medicine. Clin Trials 2008;5(3):181–93.

[36] Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy–a genomewide study. N Engl J Med 2008;359(8):789–99.

[37] Braunholtz DA, Edwards SJ, Lilford RJ. Are randomized clinical trials good for us (in the short term)? Evidence for a "trial effect". J Clin Epidemiol 2001;54(3):217–24.

[38] Amado RG, Wolf M, Peeters M, Van CE, Siena S, Freeman DJ, et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. J Clin Oncol 2008;26(10):1626–34.

[39] Hayes DF, Thor AD, Dressler LG, Weaver D, Edgerton S, Cowan D, et al. HER2 and response to paclitaxel in node-positive breast cancer. N Engl J Med 2007;357(15):1496–506.

[40] Takano T, Fukui T, Ohe Y, Tsuta K, Yamamoto S, Nokihara H, et al. EGFR mutations predict survival benefit from gefitinib in patients with advanced lung adenocarcinoma: a historical comparison of patients treated before and after gefitinib approval in Japan. J Clin Oncol 2008;26(34):5589–95.

[41] Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, Beelen K, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. Cancer Cell 2007;12(4):395–402.

[42] Iorns E, Turner NC, Elliott R, Syed N, Garrone O, Gasco M, et al. Identification of CDK10 as an important determinant of resistance to endocrine therapy for breast cancer. Cancer Cell 2008;13(2):91–104.

[43] Juul N, Szallasi Z, Eklund AC, Li Q, Burrell RA, Gerlinger M, et al. Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials. Lancet Oncol 2010;11(4):358–65.

[44] Bartlett JM, Munro AF, Dunn JA, McConkey C, Jordan S, Twelves CJ, et al. Predictive markers of anthracycline benefit: a prospectively planned analysis of the UK National Epirubicin Adjuvant Trial (NEAT/BR9601). Lancet Oncol 2010;11(3):266–74.

Part VI

Fundamental Pharmacogenomics

This page is intentionally left blank

Chapter 16

CYP450 and Ethnicity

Joseph McGraw

Metabolism Laboratory, Department of Pharmaceutical and Administrative Sciences,
Concordia University School of Pharmacy, USA

Chapter Outline

16.1 Overview **323**
16.2 Variation: Importance of Race/Ethnicity **326**
16.3 Variation: CYP450 Phenotyping **327**
16.4 CYP450 Variation in Different Racial/Ethnic Populations—CYP450 Families **329**
16.4.1 CYP 1 329
16.4.2 CYP2C8/9/19 330
16.4.3 CYP2D6 333
16.4.4 CYP3A4 and CYP3A5 333
16.4.5 Other CYP450 Isoforms 335
16.5 Future Perspectives **337**
References **337**

16.1 OVERVIEW

Racial/ethnic differences in drug response lay at the foundation of pharmacogenetics and pharmacogenomics. "Pharmacogenetics" refers to research involving heredity, drug pharmacokinetics, and pharmacologic response. It began with epidemiologic investigations of the underlying reasons for differences in response to, or inherent toxicity of, drug therapy. As with many medical studies performed in this early era, careful observation guided much of the research. Early studies implicated demographic factors such as race/ethnicity and metabolic factors such as drug clearance as important variables related to response and/or toxicity.

In the 1950s, Clayman and Hockwald et al. observed different toxicity rates in the antimalarial primaquine among African Americans versus Caucasians [1,2]. In close chronologic proximity, Hughes and Vogel et al. showed different rates of isoniazid elimination due to slow and rapid individual acetylation phenotypes [3,4]. These early studies pointed to the fact that genetic factors were likely responsible for the observed differences in drug response or toxicity among different racial/ethnic groups. In 1959, Vogel coined the term *pharmacogenetics* [5].

Pharmacologic investigation is often categorized as pharmacokinetic or pharmacodynamic in nature. Much early work acknowledged genetic drivers of pharmacologic response, but focused on pharmacokinetics (also known as metabolic studies) instead of pharmacodynamics. This is likely due, in part, to the lack of molecular biology tools for pharmacodynamic study at the time. Additionally, there was a plethora of pharmacokinetic knowledge based on robust analytical techniques already in place by the 1950s. Indeed, receptor-mediated response was difficult to assess in the early days of pharmacology while robust chemical assessment tools were commonly available. The more recent revolution in molecular biology over the last few decades has made receptor isolation and functional characterization less of a monumental task. Other more recent advances in human genetics have spurred growth in research linking human genetics and pharmacology so that now we are attempting to achieve the goal of "personalized medicine." In addition to researchers, the discipline of pharmacogenetics has garnered interest from the mainstream. The goal of the U.S. Food and Drug Administration (FDA) Critical Path Initiative 2010 is to "continue improving risk–benefit balance of approved drugs by enhancing drug product label language to include pharmacogenetics, where appropriate" [6].

Hepatic CYP450 enzymes were identified as major determinants of variability in drug metabolism early in pharmacogenetics, and their importance has not waivered [7].

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00016-5
© 2014 Elsevier Inc. All rights reserved.

Genetic analyses of CYP450s began with identification of single-nucleotide polymorphisms (SNPs) in exons; they have expanded to identify other genetic variations such as promoter region polymorphisms and gene duplications.

Pharmacokinetic differences are important determinants of variability. They were the focus of early studies and continue to be studied today. In pharmacokinetics, it is not surprising that one of the first families of oxidative enzymes characterized was the hepatic CYP450 enzyme family. Approximately 75% of commonly used drugs are CYP450 substrates, with CYP450-mediated oxidation responsible for approximately 80% of oxidative drug metabolism [8,9].

The variability of human CYP450 is complicated by the induction and inhibition of CYP450s by other drugs and xenobiotic exposures in addition to the target drug. Host variability in enzyme activity, combined with induction and inhibition, can lead to significant drug–drug interactions [10,11]. Many drug labels contain CYP450 related pharmacogenetic information. Tables are presented later in the chapter with regard to specific CYP450 isoforms and their respective substrates and/or inhibitors for which pharmacogenetic information is available in the drug package insert [12].

Hepatic CYP450s are oxidative enzymes responsible for most human oxidative xenobiotic metabolism. This metabolism is known as phase I metabolism and is often followed by conjugation of polar adducts in a process known as phase II metabolism. CYP450s participate in biosynthesis of endogenous compounds, such as steroids. However, these biosynthetic CYP450s share a specific relationship with their respective substrates and do not participate in xenobiotic metabolism. CYP450 isoforms with sequence homologies of 40% or more are categorized into families, with subfamilies based on homologies of 55% or more. Human CYP450 isoforms fall into 18 family groups with 44 subfamilies. There exist 57 sequenced human genes and 58 pseudogenes [13]. Pseudogenes are CYP450 genes that lack functionality. Xenobiotic-metabolizing CYP450 isoforms belong almost exclusively to CYP families 1, 2, and 3. They are more promiscuous in their substrate specificity compared to the other human CYPs that have specific endogenous ligands.

The primary xenobiotic-metabolizing hepatic CYP450 isoforms are CYP1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1, 3A4, and 3A5. These isoforms account for approximately 90% of CYP450-mediated drug metabolism [14]. CYP450s are heme-containing enzymes that catalyze different reactions, most commonly hydroxylation of their substrates. CYP450-mediated metabolism requires the cofactor NADPH and CYP450 reductase, and cytochrome b5 is required for maximal activity for several hepatic CYP450s. The CYP450 catalytic steps are still under investigation but are believed to proceed as follows:

1. Substrate binding near the heme iron in the ferric form.
2. Substrate binding displaces water from the heme iron.
3. The heme is reduced to its ferrous form via an electron transferred from CYP450 reductase.
4. Oxygen binds the heme iron.
5. A second electron is transferred to the bound oxygen.
6. An unstable peroxy intermediate forms.
7. A water molecule is released, generating compound I, which is a reactive iron (IV) oxo (or ferryl) radical that abstracts a proton and electron from the substrate.
8. An iron (IV) hydroxide complex is formed (known as Compound II).
9. Compound II rapidly recombines with the substrate, yielding the hydroxylated product [15]. After fully cycling through the oxidative reaction, there is a net input of one oxygen molecule, one substrate molecule, two protons, and two electrons, accompanied by the net release of a hydroxylated product and one water molecule (Figure 16.1).

Although CYP450s are expressed in every tissue of the body, hepatic CYP450s are the dominant CYP450 in regard to impact on the systemic pharmacokinetics of drugs and xenobiotics. CYP450 may be inducible, constitutively expressed, or both, and although much of its transcriptional apparatus has been delineated it is still an active area of research. The transcriptional apparatus of CYP450 varies by tissue along with expression. In the last few decades, research has shown that intestinal CYP450 is also a major determinant of xenobiotic pharmacokinetics, although to a lesser degree compared to hepatic enzymes.

Genetic polymorphisms in human xenobiotic-metabolizing CYP450 isoforms provide a major source of variability in CYP450 activity. Polymorphisms may exist that result in alteration of the physical CYP450's protein structure, gene copy number, or inducibility. An alteration in CYP450 protein structure may alter the protein's substrate affinity and catalytic activity. Often, polymorphisms result in alteration in functional phenotype. CYP450 2D6 is a CYP450 for which an individual may harbor multiple copies of CYP450 alleles. If the allele copies code for functional enzymes, multiple copies confer an enhanced metabolic phenotype. Newer polymorphisms have been identified in the promoter region of various CYP450 isoforms which alter their inducibility. They may enhance or interrupt transcription-factor binding, resulting in either enhanced or diminished inducibility, respectively (Figure 16.2).

Drug–drug interactions provide another source of variability through CYP450. Inhibition or induction of CYP450-mediated metabolism may occur because of concomitant exposure to multiple drugs or xenobiotics. Several mechanisms of CYP450 inhibition exist, but they all result in elevated exposure to drugs that would normally have been metabolized by the inhibited CYP450 isoform. Figure 16.3 illustrates the most common CYP450 inhibition mechanism via competitive inhibition. In this scenario, another

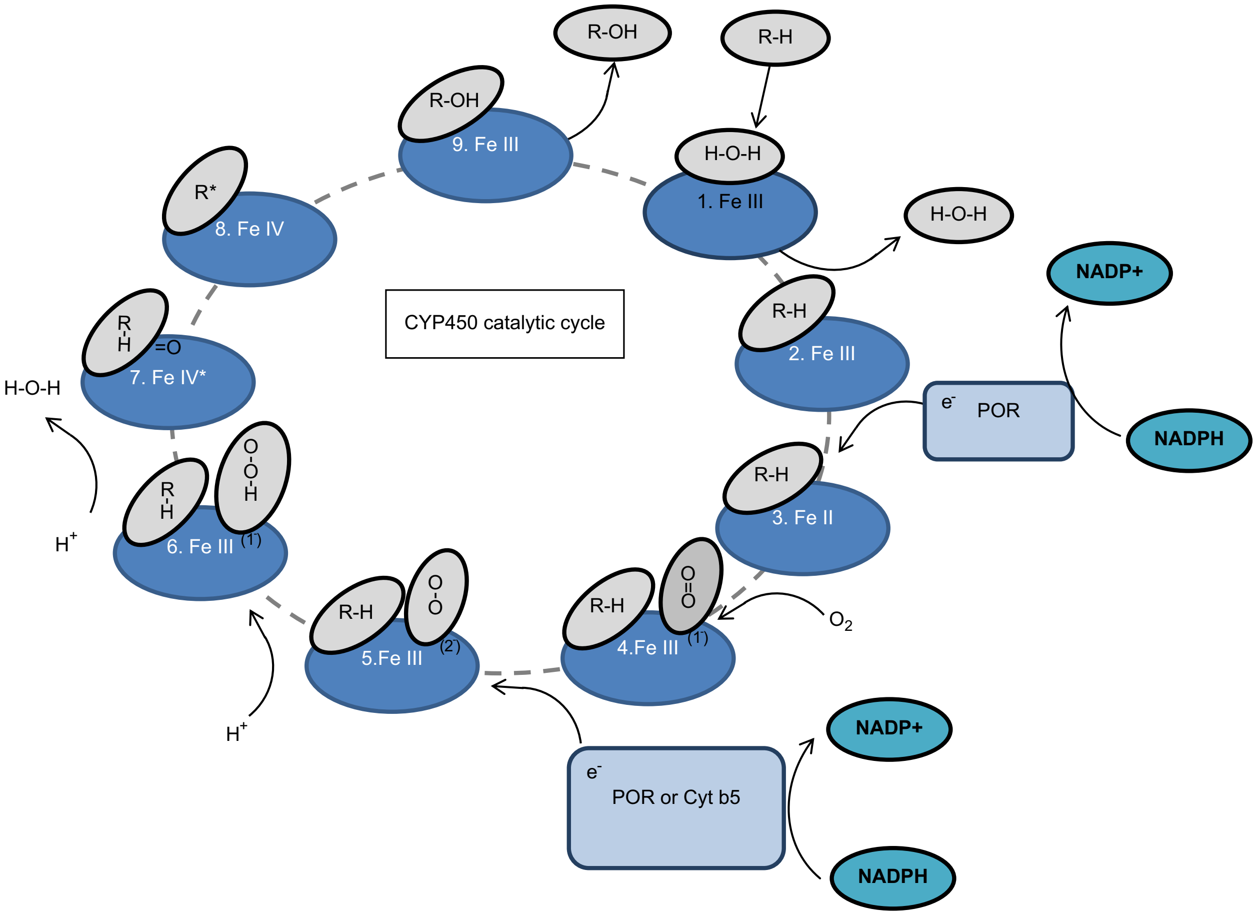

R-OH
R-H
R-OH
9. Fe III
H-O-H
1. Fe III
H-O-H
R*
8. Fe IV
CYP450 catalytic cycle
R-H
2. Fe III
NADP+
R-H
=O
7. Fe IV*
H-O-H
e-
POR
NADPH
O-O-H
R-H
6. Fe III (1)
R-H
3. Fe II
H+
R-H
O-O
5.Fe III (2)
R-H
O=O
4.Fe III (1)
O2
H+
e-
POR or Cyt b5
NADP+
NADPH

FIGURE 16.1 CYP450 catalytic cycle.

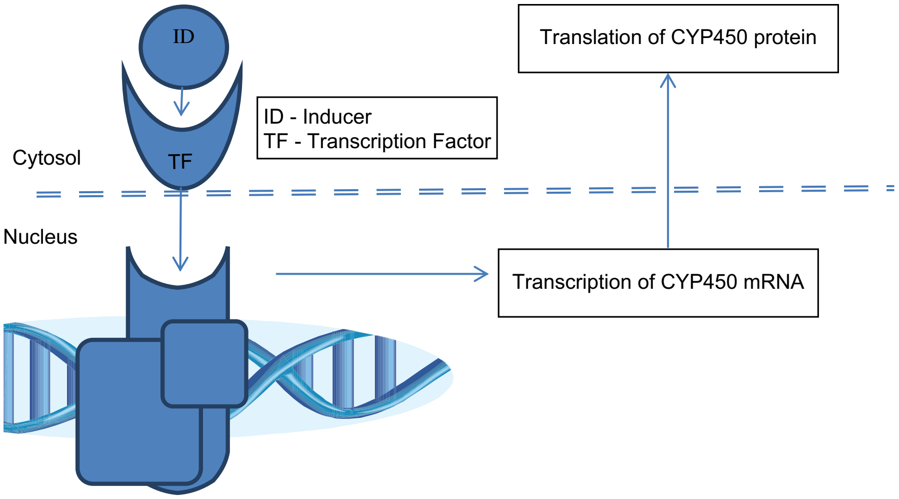

ID
Translation of CYP450 protein
ID - Inducer
TF - Transcription Factor
Cytosol
TF
Nucleus
Transcription of CYP450 mRNA

FIGURE 16.2 CYP450 induction.

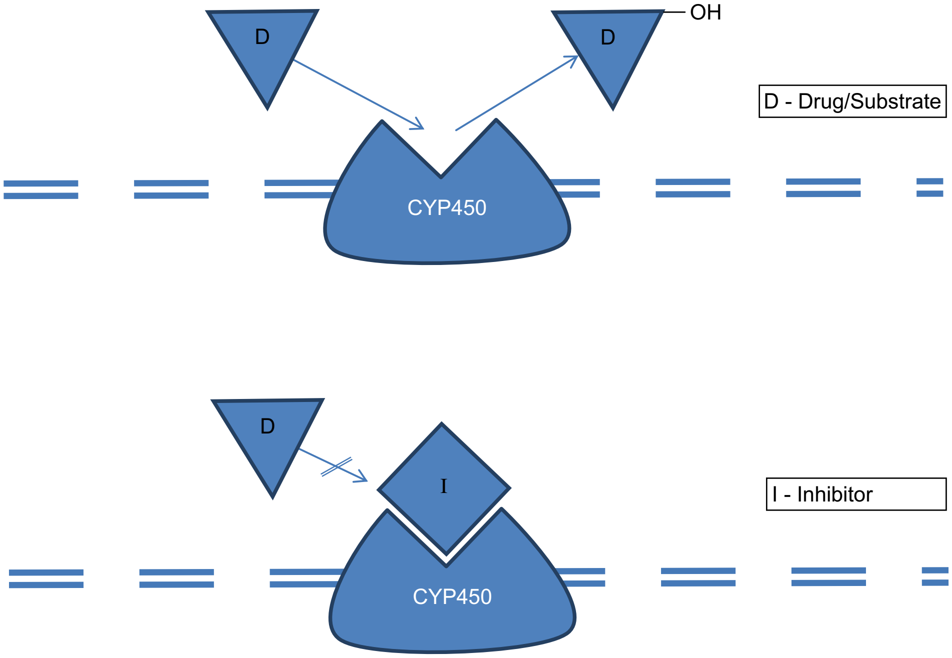


FIGURE 16.3 CYP450 inhibition.

drug prevents a CYP450 substrate from being metabolized by occupying the CYP450 isoform active site and blocking substrate access. Effects due to CYP450 inhibition may be further complicated by shunting the bulk of substrate metabolism toward another less efficient CYP450 isoform. Although such scenarios are currently being experienced by patients today, they likely go unnoticed because of their complexity. These types of scenarios will also increase in the future as the prevalence of polypharmacy increases.

Of the approximately 75% of drugs that are CYP450 substrates, about 40% are metabolized by highly polymorphic CYP450 isoforms [9,16]. As drug manufacturers continue to develop new drugs, the percentage contribution of polymorphic enzymes involved in metabolism is expected to decline. CYP isoform-mediated metabolism characterization is a prerequisite for approval, and most manufacturers avoid new chemical entities that are polymorphic CYP450 substrates.

The reported percentage of CYP450 in liver varies likely because of population differences such as racial/ethnic makeup. Also, researchers have reported different specificities of antibodies used in Western blots [17,18]. Many literature reports about CYP450 activity are quite different from one another. Often apoprotein variability is reported in terms of fold activity for many CYPs, such as CYP3A4; however, *in vivo* activity is probably less variable. Even the highly polymorphic CYP2C19 showed only 21-fold *in vivo* interindividual variability [19]. A report by Galetin et al. showed CYP3A4 that fold activity differences in healthy individuals was only 4–10-fold compared to a 40–50-fold variability in apoprotein expression shown in other studies [18]. Table 16.1 summarizes the percentage contribution of hepatic CYP isoforms to total CYP450, the percentage of drugs metabolized by each isoform, and the *in vivo* (when reported) fold variability in isoform activity.

16.2 VARIATION: IMPORTANCE OF RACE/ETHNICITY

Race is a social construct. However, individuals tend to choose mates and reproduce within perceived racial/ethnic groups, and thus allele frequencies between groups may differ. In addition, haplotype blocks may vary by racial ethnic group because of differences in recombination sites. Racial/ethnic stratification may be necessary for the specific population genetic statistical approaches used by researchers. Hardy-Weinberg Equilibrium (HWE) calculations are a common way to assess whether a genotype is distributed as expected in a cohort. Genes not in HWE may be associated with a specific aberrant phenotype. Since HWE calculations assume random mating of individuals in a group, calculations of HWE require stratification of the different racial/ethnic groups being studied. These groups are often stratified in investigations where racial/ethnic differences might confound the disease–genetic relationship. However, self-reporting of race/ethnicity can be inaccurate in terms of genetic composition. This has led some investigators to identify ancestral genotype when working with genetically heterogeneous groups [24]. It should be kept in mind that the total impact of race/ethnicity on overall genetic variation is quite minimal. Only 5–15% of genetic variation is attributable to ancestral populations living on different continents [25].

TABLE 16.1 Variability in and Proportions of CYP450 in Human Liver

CYP	Isoform Contribution to Total Liver CYP450 (%)	Drugs Metabolized by CYP450 Isoform (%)	Isoform Fold Interindividual Variability
1A2	8–13	10	40–130
2A6	4–13	3	20–>100
2B6	2–10	4	20–250
2C8	7–12	5	2–9
2C9	10–20	10–20	3–27
2C19	4–13	5–13	21–28
2D6	1–2	15–25	>1000
2E1	2–9	3	20
3A4	20–50	36–50	40–50 but 4–10-fold in healthy populations
3A5	0.2–2	Expected to be similar to 3A4	ND

ND = not reported.

Source: Based on data reported in a number of references. See [14,16,18–23].

Regardless of type, investigators usually set up studies in similar ways. They attempt to create an "all things being equal" scenario except for the exposure, intervention, or genetic differences of interest. Since the genetic variation attributable to racial/ethnic groups is minimal, the genetic differences responsible for phenotypic differences among these groups become easier to identify. Small genetic differences may be harnessed to identify loci and genes of interest. Several researchers have used these differences in disease frequency to identify disease gene loci involved in complex diseases such as cancer [26].

Not surprisingly, genetic variability among racial/ethnic groups is not equal for all genes. Genes linked to environmental responses are more likely to vary by geography and thus differ by race/ethnicity. A major driver of such genetic variation is selection pressure. Genes that confer a survival advantage will be selected in a population residing in a particular environment. Xenobiotic-metabolizing CYP450s fall into this category because different CYP450 phenotypes can confer a survival advantage [27]. An example is a CYP450 metabolic phenotype that confers resistance to toxicity from ingestion of a plant endemic to a certain region [28]. Another example is a CYP450 phenotype that resists bioactivation of a fungal carcinogen endemic to tropical regions [29]. The observed racial/ethnic differences in both genotype and phenotype for xenobiotic-metabolizing CYP450s are likely due to ancestral environment–related positive selection.

In the case of xenobiotic-metabolizing CYP450 genes, there are many examples of allele and haplotype frequency differences in racial/ethnic groups. Many reports have thoroughly described such observations, which lie at the foundation of pharmacogenetics and pharmacogenomics. One recent example is given by the Pedersen et al., who investigated the frequency of the newly discovered CYP2C19 hyperfunctional allele CYP2C19*17 in Scandinavians and Ethiopians [30]. CYP2C19*17 appears to have higher frequency in Caucasians and some Eastern African groups. However, it was also found to be in linkage disequilibrium with the other fully functional CYP2C isoform alleles for 2C8 and 2C9. In other words, the hyperfunctional CYP2C19 allele is more likely to be found on the same chromosome as the wild-type CYP2C8 and CYP2C9 alleles. In haplotype predictions, 99.7% of Scandinavian CYP 2C19*17 carriers also possess CYP2C8*1 and CYP2C9*1 on the same haplotype. Out of 10 modeled haplotypes, the aforementioned haplotype was the second most frequent (19% frequency), following wild-type CYP2C8/9/19. The selective advantage of enhanced CYP2C function is unknown, but is surely not linked to modern drug exposures.

16.3 VARIATION: CYP450 PHENOTYPING

Assessment of human hepatic CYP450 activity involves both invasive and noninvasive methods. However, each method has inherent benefits and drawbacks. Invasive methods usually involve biopsy of hepatic tissue followed by *in vitro* reconstitution-system assays to assess drug-metabolizing capacity. Reconstituted metabolism systems used to identify CYP450 activity include harvested hepatocytes, hepatic tissue slices, or microsomal preparations from liver biopsy tissue. Besides being highly invasive, these approaches are subject to other problems in interpretation. Cultured hepatocytes begin to lose their ability to express CYP450 soon after being extracted. Microsomes maintain their metabolic activity for long periods of time, but require significant amounts of tissue for testing. Many reconstitution systems do not reflect variability due to other physiologic parameters that affect metabolism, such as hepatic blood flow. *In vitro*-to-*in vivo* extrapolation models are emerging that incorporate *in vitro* data on drug-metabolism reconstitution systems into highly predictive pharmacokinetic models. These models also incorporate physiologically based pharmacokinetic model parameters, such as hepatic blood flow, thus compensating for some of the limitations of reconstitution systems.

Substrate probes are drugs that undergo hepatic CYP450–mediated metabolism selectively via one CYP450 isoform. Once the probe is administered, elimination of specific metabolites may be monitored to characterize CYP isoform activity. Substrate probes can be used in both *in vitro* and *in vivo* reconstitution systems.

Avoidance of the need for invasive techniques on human subjects has been achieved by the use of animals in pharmacokinetic models. However, many animal species have dramatically different CYP450 activity at the isoform level. Exceptions include primates with robust CYP450-isoform sequence homology to humans such as cynomolgus monkeys. However, even these models do not replicate human pharmacokinetics perfectly for all drugs metabolized by certain isoform combinations. Currently, pooled human hepatic microsomes are readily available from vendors such as BD Biosciences; these pools are created from large numbers of individuals (e.g., n = 150), so they are reflective of population-level metabolism. Recombinant technologies allow transfection of human genes into other cell systems, such as BD Biosciences' baculovirus-transfected insect cells. These transfections allow for investigation of metabolism at the individual isoform level, providing a valuable tool for investigations of individual isoform–mediated metabolism and its contribution to overall metabolism. Other tools for identifying individual-isoform metabolism are CYP450 isoform–specific monoclonal antibodies, drug phenotyping probes, specific inhibitors, and specific inducers.

Drug-phenotyping probes were the first tools used to identify isoform-specific metabolism *in vivo*, and they are still effective tools. Before our current knowledge of isoform-specific CYP450 metabolism, pharmacokinetic studies showed differential metabolism among individuals due to genetic differences. Since these differences were difficult to observe if the metabolic clearance was attributable to multiple enzymes, investigators suspected a specific relationship between the substrate and the liver enzyme that metabolized it. Examples include Smith, Echelbaum, as well as Kupfer and Wendlunds' discovery of variability in the respective metabolic phenotypes of debrisoquine, sparteine, and mephenytoin due to polymorphisms in drug oxidative metabolic enzymes [3]. Variable activity for both debrisoquine and sparteine would later be attributed to CYP2D6 polymorphisms, while variability in mephenytoin activity is now known to be due to polymorphisms in CYP2C19. Not surprisingly, CYP2D6 and CYP2C19 are two of the most variable CYP450 isoforms expressed in human liver. Early studies showed that drug clearance serves as a noninvasive mechanism to probe enzyme function. Hence, these drugs were among the first drug "probes" of CYP450 activity. In 2004, Zhou et al. created a list of recommended CYP450 substrate probes after reviewing the literature and recommendations from the Committee for Proprietary Medicinal Products (CPMP), the FDA, the online drug interaction table by Flockhart, the European Federation for Pharmaceutical Sciences (EUFEPS), and the American Association of Pharmaceutical Sciences (AAPS) [16].

Establishing *in vivo* metabolic phenotypes is not straightforward. There is no strict agreement on phenotype nomenclature, nor are there clearly defined phenotype definitions. Phenotype is often assessed by administering different substrate probes and reporting metabolic ratios of different metabolites in different matrices at different times, using different sampling methods (extraction and detection). For example, one study might assess CYP1A2 phenotype by administering caffeine, detecting it and its major metabolite, paraxanthine, in urine obtained 12 h after dosing, extracting the parent and metabolite using solid-phase extraction, and quantifying them using HPLC with UV detection. However, another study might administer caffeine and then identify it, along with paraxanthine and other metabolites in plasma, 6 h after dosing, using liquid/liquid extraction and detecting it using HPLC MS/MS. Table 16.2 shows the major parameters that may vary when performing *in vivo* substrate-probe CYP450-phenotyping studies.

The FDA has provided recommendations to help manufacturers with pharmacogenomic submissions accompanying new drug applications, but it does not define a systematic approach for all CYP450 and it focuses on *in vitro* approaches [32]. Two reviews propose steps for validation of probe assays and cocktail assays [33,34]. In some studies, phenotypic groups are designated by different names or the number of phenotypic categories varies. According to several researchers, the distribution of metabolic phenotype must exhibit multiple modes if variant alleles that impact activity are present. Also, a phenotype directly determined by alleles that code for enzymes with different activity (such as wild-type and loss-of-function phenotype) should result in a histogram with at least two modes (a.k.a bimodal distribution). However, one could argue that phenotype is determined by a mixture of epigenetic, environmental, and multigenic factors. In this setting, the impact of known CYP450 variant alleles on metabolic phenotype may not drive a population into a clearly defined multi-modal distribution.

TABLE 16.2 Major Parameters in Phenotype Measurement

Substrate Probe	Metabolite(s) Measured	Matrix	Time (h)	Extraction Method	Sample Detection Method
Multiple probes for each CYP450 Isoform (e.g., caffeine or phenacetin for CYP1A2)	Most abundant metabolite (e.g., paraxanthine for caffeine)	Urine	Often varies by study (e.g., 4, 6, 8, 10, 12)	Solid-phase	HPLC with different detection (ex. UV)
	Several metabolites (e.g. paraxanthine, and 1,7 dimethyluracil)	Saliva		Liquid/liquid extraction	GC with different detection (ex. FID)
	All identified metabolites	Serum		On-column cleanup	GC MS or GC MS/MS
		Plasma		Direct injection	HPLC MS or HPLC MS/MS

Source: Based on data reported in Carillo et al. [31].

Many studies performed on the highly variable CYP2D6 and CYP2C19 describe four different phenotypic groups. However, it must be kept in mind that these groups are predicted genotypes based on known polymorphic alleles with altered function. The currently used names for the phenotypic groups for CYP2D6 and CYP2C19 are: Poor Metabolizer, Intermediate Metabolizer, Extensive Metabolizer, and Rapid Metabolizer. Other genotype category names used include Intermediate Metabolizer, Poor/Rapid Heterozygotes, Rapid Heterozygotes, and Ultra-Rapid Metabolizer.

Pharmacokinetic studies with various substrates have revealed either an additive or dominant mechanism of inheritance for the metabolizer traits. Phenotype is thus determined by the sum of allele functions or by carriage of an allele with profoundly altered activity, such as a lack of non-functional allele or a hyperfunctional allele. Initial pharmacokinetic and platelet-response studies with clopidogrel suggested that mixed carriers of the CYP2C19*2 (a loss of function allele) and CYP2C19*17 allele (a gain of function allele) have a metabolizer phenotype similar to homozygous wild type (*1) carriers [35]. However, later studies show that the phenotype manifests as lower activity than homozygous wild type carriers [36]. Figure 16.4 gives an example of the frequency of different genotypes for CYP2C19 along with respective predicted metabolizer phenotype for each. Readers should keep in mind that there is still significant variability of true phenotype in each predicted phenotype category. This is due to other factors such as induction, drug-drug interactions, or other genes that may impact pharmacokinetics such as transporters.

16.4 CYP450 VARIATION IN DIFFERENT RACIAL/ETHNIC POPULATIONS—CYP450 FAMILIES

Cytochrome P450 variations occur in different ethnic populations for most CYP450 isoforms. Clinically relevant variants are discussed to assist in identification of the impact of variants on the metabolism of specific drugs or xenobiotics when applied across many ethnic groups. Tables of drugs follow discussions of each CYP isoform. The drugs have either U.S. or European labeling comments and/or dosing recommendations because of pharmacogenetic-related impacts on drug pharmacokinetics (see Tables 16.3–16.6).

16.4.1 CYP 1

The human CYP 1 family includes CYP1A1, 1A2, and 1B1. CYP1A1 and CYP1B1 are important enzymes involved in xenobiotic metabolism but do not participate in much drug metabolism. Both are highly induced by Aryl Hydrocarbon receptor (AHR) binding ligands, such as dioxin, while also acting as metabolic enzymes for many of these substrates. CYP1A1 and CYP1B1 are expressed mostly in nonhepatic tissues which is why they are seldom studied in regards to drug metabolism [39]. The primary hepatic CYP 1 family enzyme is CYP1A2, which metabolizes approximately 10% of drugs including caffeine, ropinirole, theophylline, and tizanidine along with several important psychiatric medications including clozapine, haloperidol, olanzapine, and trazodone [40,41]. CYP1A2 is constitutively expressed, inducible, and subject to epigenetic regulation [42]. Drug inducers of CYP1A2 include albendazole, lansoprazole, omeprazole,

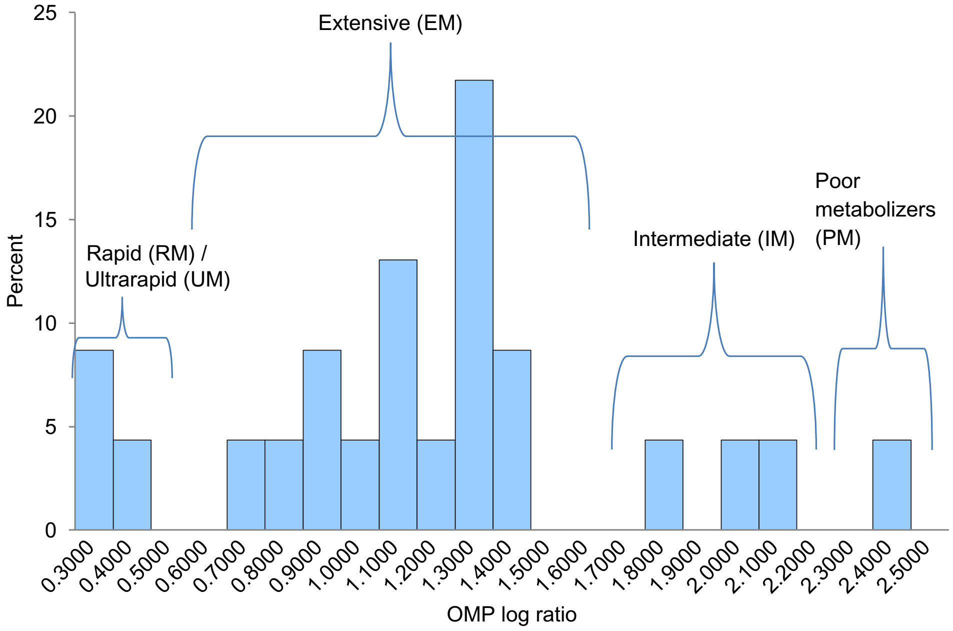


FIGURE 16.4 **Histogram of omeprazole metabolic ratios and proposed phenotype groups.** The histogram is based on a log [omeprazole]/[5-OH omeprazole] ratio measured 4 hr after a 30 mg CR omeprazole dose in saliva was extracted via molecular sieve filtration followed by HPLC MS/MS analysis (unpublished data, McGraw and Bichler 2013). CYP2C19 ultrarapid metabolizers possess two gain-of-function alleles (*17/*17); rapid heterozygotes harbor one *1 allele and one gain-of-function allele (for example, *1/*17); poor/rapid heterozygotes harbor one reduced-function and one gain-of-function allele (for example, *2/*17); extensive metabolizers are homozygous for the nonvariant *1 allele (*1/*1); intermediate metabolizers possess one *1 allele and one reduced-function allele (for example, *1/*2); and poor metabolizers harbor two reduced-function alleles (for example, *2/*2). *Source: Adapted from Scott et al.* [36].

TABLE 16.3 US and/or European Drugs with CYP1A2 Related Pharmacogenetic Labeling Comments or Dosing Recommendations

CYP1A2-US	
Fluvoxamine[a]	Olanzapine[a]
CYP1A2-Europe	
None	None

[a] *U.S. and/or European labeling recommendation.*

Source: Based on data reported in Whirl-Carrillo et al. [50].

and primaquine while contraceptives, fluvoxamine, and several quinolones are important inhibitors [40,41]. CYP1A2 activity varies widely (40–130-fold), due in part to its induction via several commonly consumed chemical exposures such as caffeine, charcoal barbecued meats, cigarette smoke, omeprazole, and oral contraceptives [43].

Genetics appear to be responsible for approximately 35% of variability in CYP1A2 activity [44]. CYP1A2 phenotypes are often categorized as extensive and poor while some studies report an ultrarapid phenotype. Metabolic ratio distributions have been described as bimodal or trimodal [45]. However, most studies do not report phenotypic groups beyond poor and normal (i.e., extensive) metabolizers. The frequency of poor metabolizer-phenotype status varies in different racial/ethnic groups. The two major haplotypes associated with it are known as CYP1A2*1C (rs2069514) and CYP1A2*1K (rs2069526, rs12720461, and rs762551). The CYP1A2*1C allele is commonly found in Asians (frequency ~25%) and is associated with decreased function, while the CYP1A2*1K allele (Asian frequency 0–4%) is associated with lower inducibility. Examples of poor metabolizer frequency in different populations are Australians (5%), Chinese (5%), and Japanese (14%). Overall, Asian and African populations have lower CYP1A2 activity compared to Caucasians [18]. The CYP1A2*1F allele (rs762551) appears to enhance the inducer function of CYP1A2 activity in carriers exposed to inducers such as cigarette smoke or omeprazole [46,47]. It is highly prevalent with a similar distribution across different racial/ethnic groups (frequency 60–66%) [47–49].

16.4.2 CYP2C8/9/19

Human CYP 2C-family enzymes share more than 80% sequence homology and exhibit both overlapping and

TABLE 16.4 US and/or European Drugs with CYP2C Family Related Pharmacogenetic Labeling Comments or Dosing Recommendations

CYP2C8-US			
None		None	
CYP2C8-Europe			
None		None	
CYP2C9-US			
Celecoxib[a]	Flurbiprofen[a]	Fluvoxamine[a]	Glimepiride[b]
Phenytoin[b]	Tolbutamide[b]	Warfarin[c]	
CYP2C9-Europe			
Acenocoumarol[b]	Gliclazide[b]	Glibenclamide[b]	Phenprocoumon[b]
CYP2C19-US			
Amitriptylline[b]	Carisoprodol[a]	Citalopram[c]	Clobazam[a]
Clomipramine[b]	Clopidogrel[c]	Diazepam[a]	Doxepin[b]
Drospirenone[a] & ethinyl estradiol	Escitalopram[b]	Esomeprazole[c]	Fluvoxamine[a]
Imipramine[b]	Lansoprazole[c]	Modafinil[a]	Nelfinavir[a]
Omeprazole[c]	Pantoprazole[c]	Prasugrel[a]	Rabeprazole[c]
Sertraline[b]	Voriconazole[c]		
CYP2C19-Europe			
Moclobemide[b]			

[a]*U.S. and/or European labeling recommendation.* [b]*Dosing recommendation.* [c]*Labeling and Dosing recommendations.*
Source: Based on data reported in Whirl-Carrillo et al. [50].

distinctive substrate specificity. They represent about 20% of liver CYP450 and metabolize about 25% of clinical drugs.

The least significant drug-metabolizing CYP2C isoform in terms of expression, CYP2C8, represents about 7% of total hepatic CYP450 and metabolizes around 5% of clinical drugs, including cerivastatin, several NSAIDs, paclitaxel, repaglinide, and thiazolidinediones. It is inhibited by gemfibrozil [51]. Several CYP2C8 polymorphisms believed to be clinically relevant include CYP2C8*1B, *2, *3, *4, and *5. Carriage of CYP2C8*2, a fairly prevalent allele, is believed to result clinically in impaired metabolic activity [52]. Conflicting results in the identification of CYP2C8 phenotype may be due to confounding factors such as participation of other enzymes (e.g., overlapping substrate specificity with CYP2C9) and linkage disequilibrium. Haplotype analyses may help alleviate these problems.

The CYP2C8*3 allele exhibits lower activity *in vitro*, but has been associated with higher clearance of CYP2C8 substrates *in vivo*. Rodriguez-Antona et al. showed increased clearance in the CYP2C8*3 allele containing haplotype carriers (haplotype B or haplotype D). Paclitaxel clearance was higher in haplotype B carriers, and repaglinide (a thiazolidinedione) clearance was higher in haplotype B or haplotype D carriers who were also heterozygous SLCO1B1 521T/T carriers (SLCO1B1 carriers have increased repaglinide uptake) [53]. Both drugs were metabolized to a lesser extent in the CYP2C8*1B-containing haplotype C carriers. The activity of the CYP2C8*4 allele is unknown, and the CYP2C8*5, *7, and *11 alleles are low frequency loss of function alleles.

The CYP2C9 alleles with impaired function observed clinically are 2C9*2, 2C9*3, and 2C9*5. CYP2C9 and 2C19 share 92% sequence homology [54]. CYP2C9 substrates include the angiotensin receptor antagonist losartan, the anticoagulant warfarin, the anticonvulsant phenytoin, the antidepressant fluvoxamine, the anti-inflammatories (celecoxib, diclofenac, flurbiprofen, and ibuprofen), fluvastatin,

TABLE 16.5 US and/or European Drugs with CYP 2D6 Related Pharmacogenetic Labeling Comments or Dosing Recommendations

CYP2D6-US			
Amitriptyline[c]	Aripiprazole[c]	Atomoxetine[c]	Carvedilol[c]
Citalopram[a]	Clomipramine[c]	Clozapine[c]	Codeine[c]
Desipramine[c]	Dextromethorphan[a]	Doxepin[c]	Duloxetine[b]
Flecainide[b]	Fluoxetine and Olanzapine[a]		Fluvoxamine[a]
Galantamine[a]	Geftinib[a]	Haloperidol[b]	Iloperidone[a]
Imipramine[c]	Metoprolol[c]	Mirtazapine[b]	Modafinil[a]
Nortriptyline[c]	Olanzapine[b]	Oxycodone[b]	Paroxetine[c]
Perphenazine[a]	Pimozide[a]	Propafenone[c]	Propranolol[a]
Protriptyline[a]	Quinidine[a]	Risperidone[c]	Tamoxifen[b]
Terbinafine[a]	Tetrabenazine[a]	Thioridazine[a]	Timolol[a]
Tiotropium[a]	Tolterodine[a]	Tramadol[b]	Tramadol & Acetaminophen[a]
Trimipramine[c]	Venlafaxine[c]		
CYP2D6-Europe			
Cevimeline[a]	Flupenthixol[b]	Zuclopenthixol[b]	

[a]*U.S. and/or European labeling recommendation.* [b]*Dosing recommendation.* [c]*Labeling and Dosing recommendations.*

Source: Based on data reported in Whirl-Carrillo et al. [50].

TABLE 16.6 US and/or European Drugs with CYP3A Family Related Pharmacogenetic Labeling Comments or Dosing Recommendations

CYP3A4-US			
Aripiprazole[a]	Fluvoxamine[a]	Gefitinib[a]	Nelfinavir[a]
Ticagrelor[a]			
CYP3A4-Europe			
See Above			
CYP3A5-US			
Tacrolimus[a]			
CYP3A5-Europe			
See above			

[a]*U.S. and/or European labeling recommendation.*

Source: Based on data reported in Whirl-Carrillo et al. [50].

nateglinide, and the sulfonylureas (glyburide, glibenclamide, glimepiride, and tolbutamide). Inhibitors of CYP2C9 include amiodarone, benzbromarone, bucolome, fluconazole, miconazole, and sulphaphenazole; it is induced by rifampicin [51]. Significant interest has focused on the effect of CYP2C9 polymorphism on warfarin activity and dosing.

Limdi et al. found an association between CYP2C9 variant allele carriers and lower warfarin dosing in Caucasian Americans, but not African Americans. They also found a significantly higher prevalence of variant genotypes in Caucasians versus African Americans (29.82% vs. 9.73%) despite the *5, *6, and *11 variants observed only in African Americans [55]. Scott et al. recently found a CYP2C9 allele labeled *8 with possibly lower functionality which may have a prevalence close to 9% in African Americans. They noted that addition of this allele to screening panels makes warfarin dose prediction more accurate in African Americans [56].

CYP2C19 is highly polymorphic, with at least 24 variant alleles of which many have no enzymatic activity. Important substrates for CYP2C19 include several benzodiazepines (diazepam, etizolam), the antiplatelet agent clopidogrel, proton pump inhibitors (lansoprazole, omeprazole, and rabeprazole), phenytoin, sertraline, and voriconazole. Inhibitors include ticlopidine, fluvoxamine, and voriconazole, while ritonavir, rifampicin, and rifabutin act as inducers [51]. The variants *2–*8 have been shown to be inactive [57]. They also distribute differently across racial/ethnic groups. Unlike other CYP450s, Caucasians and Africans share similar overall frequencies of poor metabolizer (PM) phenotype (1–8%), while Asians show a higher prevalence of PMs (13–23%) [21]. Phenotypes have been designated as extensive (two functional alleles), intermediate (one functional/one dysfunctional), poor (two loss-of-function alleles) and rapid (one or two gain-of-function alleles) (Table 16.7). The lack of functional CYP2C19 is a risk factor for adverse cardiovascular outcomes in patients taking clopidogrel because clopidogrel must be bioactivated by CYP2C19 for activity [35–37].

16.4.3 CYP2D6

CYP2D6 is responsible for metabolizing a number of important drugs containing amine functional groups, including members of the following psychotropic classes: anticholinergics/parasympathomimetics, antidepressants and monoamine modulating drugs (for example, serotonin 5-HT$_3$ receptor antagonists, monoamine oxidase inhibitors (MAOIs), serotonin reuptake inhibitors (SSRIs), and tricyclic antidepressants); antipsychotics (typical and atypical); opiates; and synthetic opiate derivatives. CYP2D6 also metabolizes several cardiac drugs from antiarrhythmic classes as well as beta blockers; some antifungals; and the antiestrogen tamoxifen. It is inhibited by celecoxib, cinacalcet, quinidine, several SSRIs (paroxetine and fluoxetine), and terbinafine [51].

A noteworthy attribute of CYP2D6 is that it is not inducible; rather, it is highly polymorphic with more than 100 variant alleles and ~200-fold variability in the metabolism of at least 100 drugs [58–61]. An attribute specific to CYP2D6 in comparison to other CYP450 enzymes is the presence of gene duplications that may confer an ultrarapid metabolizer phenotype. These polymorphisms are designated CYP2D6*(gene variant)XN, where XN refers to the number of gene copies. For example, CYP2D6*1X2 represents two copies of CYP2D6*1. The CYP2D6*1XN, *2XN, and *35XN alleles confer enhanced metabolic phenotype, while CYP2D6*17XN and CYP2D6*41XN show decreased activity and CYP2D6*4XN alleles show none.

Four potential CYP2D6 phenotypic subgroups exist. These groups are usually defined by the respective number of their functional alleles: ultrarapid (3), extensive (2), intermediate (1), and poor metabolizers (0). Most CYP2D6 polymorphisms result in an allele that lacks metabolic activity. However, the prevalence of poor metabolizer phenotypes varies by racial/ethnic group: Asians (~1%), Caucasians (5–10%), and Africans (0–19%) [58].

16.4.4 CYP3A4 and CYP3A5

The human CYP3A family includes CYP3A4, 3A5, 3A7, and 3A43, with CYP3A4 being the most abundant CYP450 isoform, representing up to 50% of liver CYP450. CYP3A4 represents approximately 20–50% of CYP450 content in liver and is responsible for 36–50% of drug metabolism. CYP3A5 is a closely related homologue that represents only 2% of CYP3A in Caucasian livers [9,22]. Examples of CYP3A4 substrates include many members of several drug classes, including benzodiazepines, calcium channel blockers, corticosteroids, ergot alkaloids, statins, transplant medications, and vinca alkaloids. Major inhibitors include azole antifungals, calcium channel blockers (diltiazem, verapamil), cimetidine, grapefruit juice, several macrolide antibiotics (erythromycin, clarithromycin, telithromycin, troleandomycin), nefazodone, and protease inhibitors. Inducers include rifampicin, rifabutin, phenobarbital, phenytoin, carbamazepine, efavirenz, and St. John's Wort [41].

Intestinal CYP3A4 also makes a major contribution to the metabolism of orally administered drugs because it contributes to first-pass metabolism [8]. Constitutive CYP3A4 variability is estimated at about 5-fold; however, illness, inhibition, and induction-related interactions can enhance variability up to 400-fold [8,18]. Other estimates across different populations project 40–50-fold variability [16]. Some estimates characterize variability in terms of apoprotein expression, which may overestimate *in vivo* variability for healthy subjects not experiencing drug

TABLE 16.7 CYP2C19 - Example of Genotypes, Frequency, and Predicted Phenotype

			CYP 2C19 Genotype		
Major Phenotypic Categories	Metabolizer Phenotype	Prevalence Range (%)	Normal Allele (*1)	Loss of Function (*2, *3, …)	Gain of Function (*17)
Poor	Poor	7(±8)		X,X	
Intermediate	Intermediate	34(±17)	X	X	
	Poor/Rapid Heterozygotes	4(±4)		X	X
Extensive	Extensive	51(±15)	X,X		
Rapid	Rapid Heterozygotes	18(±17)	X		X
	Ultra-rapid	2(±4)			X,X

Source: Based on data reported in the following references [30,36–38].
Note: 2C19 East Asian prevalence of poor metabolizer phenotype may be as high as 24%.

interactions. Population CYP3A4 metabolic phenotypes follow a unimodal distribution in most studies, which has been presented as evidence that CYP3A4 does not have a genetically determined poor or rapid metabolizer phenotype [45]. However, genetic contributions to variability may be obscured when other factors, such as enzyme induction via inducer exposure, are accounted for.

CYP3A4 and CYP3A5 share significant sequence homology and almost identical substrate specificity, but exhibit somewhat differing metabolic rates. CYP3A enzymes have a large active site and accommodate a great variety of substrates. Because of the polygomous nature of the active sites, substrates, including substrate probes, are often metabolized by both CYP3A4 and CYP3A5, but a few substrate probes appear to undergo selective metabolism. Erythromycin and quinidine are more specific CYP3A4 probes, while CYP3A5 preferentially metabolizes alprazolam alpha-hydroxylation, tacrolimus, and vincristine [62]. Caucasians exhibit lower CYP3A5 activity as a group because of a high prevalence of the splice variant, CYP3A5*3 [62], which is a highly prevalent polymorphism in all ethnic groups and causes the absence of functional CYP3A5 protein. CYP3A5*3 prevalence estimates in different groups are as follows: African Americans (32%), Caucasians (90–93%), East Asians (73%), Hispanics (65%), and South Asians (60%) [63–65].

Caucasians show enhanced CYP3A4-mediated metabolism in comparison to those of African ancestry [66]. Conversely, most Caucasians carry the loss-of-function CYP3A5*3 polymorphism. In most individuals, CYP3A5 has a minimal impact, as it represents only 2% of hepatic CYP450 in Caucasians. However, it may represent up to 50% of hepatic and intestinal CYP450 in some individuals [67]. A proposed cause for the differences in CYP450 activity may be the differential activity of CYP450 cofactors such as NADPH-dependent CYP450 reductase and cytochrome b5 (see Figure 16.1).

These cofactors are variable in different groups and may contribute to different CYP3A4 and CYP3A5 metabolic activities [68–70]. However, Elens et al. found carriage of the CYP450 oxidoreductase variant POR*28 was associated with lower *in vivo* CYP3A5 activity but had no impact on CYP3A4 activity [63]. They assessed CYP3A4/5 activity by administering midazolam, a nonspecific CYP3A4/5 probe, while assessing CYP3A4 specific activity with erythromycin, a specific CYP3A4 probe. They concluded that differences in CYP3A-mediated metabolism in Caucasians versus other racial/ethnic groups are not likely explained via CYP3A5-mediated metabolism or cofactor variability. Another major source of variability is interaction between CYP3A4/5 and drug transporters. Ogasawar et al. found that CYP3A5 expressors that are also in the MRP2 (transporter) high-activity group showed up to 2.3-fold lower dose–normalized trough tacrolimus concentrations. However, the researchers found no relationship with CYP3A4 variants [71].

As with most proteins and enzymes studied in pharmacogenetics, new functionally relevant polymorphisms and haplotypes of CYP450 enzymes are being identified. CYP3A4*1B is a CYP3A4 polymorphism that may alter CYP3A4 expression. It is an A→G transition at the −293 position in the gene promoter region that is highly prevalent in African Americans but not in Caucasians. The

CYP3A4*1B prevalence among different ethnic groups is Africans (76%), Caucasians (2–9.6%), Chinese and Taiwanese (0%), Hispanic Americans (9.3–11%), and African Americans (35–67%). Although it may provide much needed answers, the clinical impact of CYP3A4*1B is still under debate [72]. Other, yet to be discovered, candidate polymorphisms may explain differential CYP450 activity in different racial/ethnic groups as they are identified. A newly identified SNP, CYP3A4*20, results in a premature stop codon, thus coding for a truncated CYP3A with no activity [22]. This SNP has a relatively high prevalence among differing ethnic groups: Caucasian (6%), African American (26%), and Chinese (22%). The CYP3A4*22 variant allele is an intron 6 SNP found to influence RNA expression and statin, tacrolimus, and cyclosporin dose requirements [63,64,72]. Several other novel variants were also recently identified in a South African cohort, including CYP3A4*24, which is predicted to affect function; it was found at prevalences of 10.3%, 3.1%, and 3.2%, respectively, in Khoisan, Xhosa, and mixed-ancestry individuals [72]. More research is necessary before a clear understanding of the mechanism behind racial/ethnic differences in CYP3A4/5 is achieved. In addition to racial/ethnic metabolism differences, it should be noted that women have higher CYP3A activity than men, but this is offset by mens' larger hepatic mass [73,74].

16.4.5 Other CYP450 Isoforms

The CYP 2 gene cluster encodes several different CYP450 isoforms (CYP2A6, 2A7, 2A13, 2B6, 2F1, and 2S1) that are not often the focus of drug metabolism studies. Several factors explain this: they are minor contributors to hepatic CYP450 content, they have few known drug substrates, and/or they are expressed only in extrahepatic tissues. Although these human xenobiotic CYP450 enzymes are not highly expressed in liver, most are distributed in other tissues throughout the body. Pavek and Dvorak provide an excellent review of the distribution of these enzymes in extra-hepatic tissues [75]. Although CYP450s are not major sources of drug metabolism, they are involved in metabolism of other xenobiotics. Phenotypic differences in xenobiotic metabolism and differential metabolism of environmental xenobiotics, such as occupational and environmental contaminants, have also been reported among different racial/ethnic groups. Research regarding extrahepatic CYP450 metabolism has led to the belief that these isoforms may play a role in susceptibility and risk for diseases such as cancer [39].

CYP2A6 is an important human xenobiotic-metabolizing enzyme. However, it metabolizes only ~3% of drugs and is not the primary metabolic enzyme for most of these, likely because of its small active site and low overall hepatic expression. CYP2A6 is primarily expressed in liver (1–10% of liver CYP450) and nasal mucosa. It bioactivates several known procarcinogens such as aflatoxin B1 and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (i.e., nicotine-derived nitrosamine ketone, NNK), while also acting as the major CYP isoform responsible for nicotine and tegafur metabolism. CYP2A6 inhibitors include isoniazid and methoxsalen as well as several endogenous catecholamines and steroids, while it is induced by phenobarbital and rifampicin [76]. CYP2A6*1A and *1B alleles are the most common CYP2A6 variants. The prevalence of CYP2A6*1A is 40% in Asians and 66% in Caucasians, whereas CYP2A6*1B is found in 40% of Asians and 30% of Caucasians [77,78].

More recently, four new dysfunctional alleles (*4G, *4H, *1B4, and *1L) have been discovered in individuals of African ancestry, with a combined prevalence of 7.3% [79,80]. African ancestry–specific polymorphisms are important because they help to explain the lower nicotine clearance observed in those of African descent when compared to those of Caucasian descent [81–83]. Nicotine metabolism differences are also important because African Americans suffer higher rates of smoking-related diseases [80]. As previously mentioned, CYP2A6 is important in the metabolism of environmental xenobiotics. Polychlorinated biphenyls (PCBs) are environmental contaminants that were banned from use in the 1970s but are still found in human tissues today. CYP2A6 was found to specifically metabolize one PCB congener, known as PCB congener 101, which is highly prevalent in the environment but should be rapidly metabolized in most humans [84]. PCB 101 was found at higher serum concentrations in African American women reporting exclusive African ancestry versus African American women of mixed race/ethnicity [85]. It was believed that higher PCB 101 levels may have been a marker of poor metabolism linked to those of African descent. This may be an environmental correlate to the lower nicotine metabolism in carriers of loss-of-function CYP2A6 polymorphisms.

CYP2B6 is highly polymorphic and highly inducible (20–250-fold variation), with 53 allelic variants described thus far [13,86]. Basal CYP2B6 expression in human liver is low, but is highly induced with relatively low hepatic expression compared to other CYP isoforms (~1–10% of liver CYP450) [87]. CYP2B6 variability among racial/ethnic groups is noteworthy for several reasons. Its phenotypic activity does not appear to vary by racial/ethnic group in men but is 3.6- and 5.0-fold higher in Hispanic women compared to Caucasian and African-American women, respectively [88]. CYP450 allele frequencies also vary highly among large racial/ethnic groupings such as Asians. For this reason, it is important to be more specific in terms of racial/ethnic stratification regarding it.

Important substrates for CYP2B6 include aflatoxin, bupropion, chlorpyrifos, cyclophosphamide, efavirenz, nevirapine, and nicotine [86,89]. It is inhibited by clopidogrel, mifepristone (RU486), selegiline, methadone, tamoxifen, ticlopidine, and thioTEPA, while it is induced by artemisinin antimalarials, carbamazepine, cyclophosphamide, efavirenz, hyperforin, metamizole, N,N-diethyl-m-toluamide (DEET), phenytoin, phenobarbital, rifampicin, ritonavir, and statins [87]. Several CYP2B6 SNPs may be present on one allele because of its high rate of polymorphism. The most common SNP present in multiple alleles is c.516G→T; this SNP is responsible for a loss-of-function phenotype and is often found in combination with c.785A→G [90]. As mentioned previously, there is a large amount of heterogeneity in the frequency of the c.516G→T allele in different Asian populations: Han Chinese (21%), Hong Kong Chinese (43%), Indian (39%), Japanese (14–20%), Korean (15%), Southern Chinese (35%), Taiwanese (14%), Thai (32%), Uygur Chinese (28%), and Vietnamese (27%) [91]. The activity phenotype of many CYP2B6 variants has not been characterized, and some variants appear to behave differently toward different substrates. CYP2B6*6 is a common variant (15–60% frequency) that confers significantly reduced activity *in vivo* attributable to aberrant splicing [92]. Many loss-of-function alleles have been identified; they include CYP2B6*2, *6, *7, *11, *15, *16, *18, *26, *27, and *28. Dysfunctional alleles have been shown to be present in more than 45% of individuals, but this figure varies widely by population. CYP2B6*18 is a loss-of-function allele prevalent in those of African ancestry (4–12%). A variant CYP2B6 genotype has been linked to altered pharmacokinetics of bupropion, cyclophosphamide, efavirenz, methadone, and neviripine [87].

CYP2E1 is a low-abundance hepatic CYP450 isoform (~5% hepatic CYP450) for which only two polymorphisms have been associated with altered *in vivo* activity [57,93,94]. Drugs metabolized primarily by CYP2E1 include chlorzoxazone, dacarbazine, enflurane, ethanol, halothane, isoniazid, sevoflurane, and theophylline; CYP2E1 is involved in toxicity of acetaminophen, ethanol, and styrene [95,96]. The CYP2E1*1D allele is associated with increased 2E1 activity in the presence of chronic ethanol ingestion and obesity [97]. CYP2E1*6 allele carriers who were alcoholics had significantly lower CYP2E1 activity [98]. CYP2E1 is located in the endoplasmic reticulum membrane, as are all CYP450s, but it is also located in the mitochondrial membrane. Induction of mitochondrial CYP2E1 increases reactive oxygen species generation and is thought to participate in the toxic effects of ethanol [99].

Table 16.8 lists helpful websites that provide more information regarding CYP450 enzymes, drug interactions, and

TABLE 16.8 CYP450, Drug Interaction, and Pharmacogenetics Websites

CYP450 Websites	Topics
www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/DrugInteractionsLabeling/ucm080499.htm#background	Drug development and interactions
http://Pubchem.ncbi.nlm.nih.gov	Drug information and interactions
http://drnelson.uthsc.edu/CytochromeP450.html	General
http://cpd.ibmh.msk.su/	General
www.icgeb.org/~p450srv/	General
http://bioinformatics.charite.de/supercyp/	Interactions and general
www.medicine.iupui.edu/Flockhart/table.htm	Interactions
www.umm.edu/adam/drug_checker.htm	Interactions
www.drug-interactions.com	Interactions
www.druginteractioninfo.org/	Interactions
www.interspeciesinfo.com/Interspecies	Interspecies metabolism
http://catalog2.corning.com/Lifesciences/media/pdf/High-Low_P450_Single_Donor_HLM_Panel_031_5_1_13.pdf	Metabolism
www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/DrugInteractionsLabeling/ucm093664.htm	Metabolism
www.cypalleles.ki.se/	Pharmacogenomics
http://pharmgkb.org/	Pharmacogenomics
www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm	Pharmacogenetic information in drug package inserts
Subscription Information Sources	
www.themedicalletter.com/adi	Interactions
www.lexi.com/institutions/products/pda/lexi-complete/	Drug information/Pharmacogenomics database

CYP450 pharmacogenetics. The listed descriptions give a brief idea of content, but most of the websites offer in-depth information beyond this.

16.5 FUTURE PERSPECTIVES

As mentioned previously, CYP450s are the primary phase I enzymes responsible for metabolizing approximately 75% of all drugs. Most of this metabolism is mediated by the xenobiotic-metabolizing, hepatically expressed CYP450 isoforms CYP1A2, CYP2C8/9/19, CYP2D6, CYP2E1, and CYP3A4/5. Early researchers identified differences in CYP450-mediated metabolism among racial/ethnic groups. Identification of the genetic determinants of these differences spawned the field of pharmacogenetics.

Today we continue to identify the genetic and environmental determinants of differences in drug response at the individual and group levels. Although race is a social construct, racial stratification is often a necessity in population genetic studies involving CYP450 genes. The contribution of interethnic genetic variation is minimal compared to interindividual variation. However, unstable genes linked to environmental circumstances, such as CYP450, vary dramatically among different racial/ethnic groups. Many examples of racial/ethnic differences in CYP450 allele frequencies and haplotype blocks have been identified. Not surprisingly, metabolic phenotype prevalences vary by race/ethnicity for most of the important CYP450 isoforms.

The wealth of pharmacogenetic data accumulated thus far regarding the aforementioned CYP450 isoforms has led to many drug-labeling notices and/or dosing recommendations for individuals with variant metabolic phenotypes who may be taking specific drugs. A barrier to their use is the lack of information regarding individual patients' metabolic phenotype. Scott et al. have provided an up-to-date list of providers of CYP2C19 genotyping, and the number of vendors offering such services is likely to increase [36]. Indeed, whole exome sequencing should become fairly common with an ever growing number of next-generation sequencing providers. However, in the case of CYP450 phenotype, exon sequencing may not provide complete answers because new clinically significant polymorphisms in nonexon regions such as the promoter region are being discovered.

Increased use of haplotype approaches in pharmacogenetics research should also prove to be beneficial. This is especially true for CYP450 isoforms that share loci, among them the CYP2C family. As more data are gathered, and as knowledge of individual pharmacogenetic information (e.g., metabolic phenotype) becomes more widespread, the goal of personalized medicine will become a reality. Enhanced knowledge of drug- and patient-specific pharmacogenetic data will allow prescribers to maximize drug efficacy for their patients while minimizing the risks of toxicity and side effects.

REFERENCES

[1] Clayman CB, Arnold J, Hockwald RS, Yount Jr EH, Edgcomb JH, Alving AS. Toxicity of primaquine in Caucasians. J Am Med Assoc 1952;149(17):1563–8.

[2] Hockwald RS, Arnold J, Clayman CB, Alving AS. Toxicity of primaquine in Negroes. J Am Med Assoc 1952;149(17):1568–70.

[3] Meyer UA. Pharmacogenetics – five decades of therapeutic lessons from genetic diversity. Nat Rev Genet. 2004;5(9):669–76.

[4] Hughes HB. On the metabolic fate of isoniazid. J Pharmacol Exp Ther 1953;109(4):444–52.

[5] Vogel F. Moderne probleme der humangenetik. Ergebn Inn Med Kinderheilkd 1959;12:52–125.

[6] FDA U. The critical path initiative: report on key achievements in 2009: US FDA, <www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/UCM221651.pdf>; 2010 [cited 11.10.10].

[7] Zdanowicz. Concepts in pharmacogenomics. American Society of Health-System Pharmacists; 2010.

[8] Wilkinson GR. Drug metabolism and variability among patients in drug response. N Engl J Med 2005;352(21):2211–21.

[9] Guengerich FP, Rendic S. Update information on drug metabolism systems–2009, part I. Curr Drug Metab 2010;11(1):1–3.

[10] Murray M. Role of CYP pharmacogenetics and drug-drug interactions in the efficacy and safety of atypical and other antipsychotic agents. J Pharm Pharmacol 2006;58(7):871–85.

[11] Preissner S, Kroll K, Dunkel M, Senger C, Goldsobel G, Kuzman D, et al. SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. Nucleic Acids Res 2009;38(database issue):D237–43.

[12] FDA U. Table of valid genomic biomarkers in the context of approved drug labels: US FDA, <www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm>; 2010 [cited 11.10.10].

[13] Nelson D. The cytochrome P450 homepage: Nelson, DR, <http://drnelson.uthsc.edu/human.P450.table.html>; 2009 [updated14.08.09; cited 29.08.10].

[14] Rendic S. Summary of information on human CYP enzymes: human P450 metabolism data. Drug Metab Rev 2002;34(1–2):83–448.

[15] Krest CM, Onderko EL, Yosca TH, Calixto JC, Karp RF, Livada J, et al. Reactive intermediates in cytochrome P450 catalysis. J Biol Chem 2013.

[16] Ingelman-Sundberg M. Human drug metabolising cytochrome P450 enzymes: properties and polymorphisms. Naunyn Schmiedebergs Arch Pharmacol 2004;369(1):89–104.

[17] Hustert E, Haberl M, Burk O, Wolbold R, He YQ, Klein K, et al. The genetic determinants of the CYP3A5 polymorphism. Pharmacogenetics 2001;11(9):773–9.

[18] Galetin A, Brown C, Hallifax D, Ito K, Houston JB. Utility of recombinant enzyme kinetics in prediction of human clearance: impact of variability, CYP3A5, and CYP2C19 on CYP3A4 probe substrates. Drug Metab Dispos 2004;32(12):1411–20.

[19] Koukouritaki SB, Manro JR, Marsh SA, Stevens JC, Rettie AE, McCarver DG, et al. Developmental expression of human hepatic CYP2C9 and CYP2C19. J Pharmacol Exp Ther 2004;308(3):965–74.

[20] Hu L, Zhuo W, He YJ, Zhou HH, Fan L. Pharmacogenetics of P450 oxidoreductase: implications in drug metabolism and therapy. Pharmacogenet Genomics 2012;22(11):812–9.
[21] Zhou SF, Liu JP, Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. Drug Metab Rev 2009;41(2):89–295.
[22] Westlind-Johnsson A, Hermann R, Huennemeyer A, Hauns B, Lahu G, Nassr N, et al. Identification and characterization of CYP3A4*20, a novel rare CYP3A4 allele without functional activity. Clin Pharmacol Ther 2006;79(4):339–49.
[23] Shimada T, Yamazaki H, Mimura M, Inui Y, Guengerich FP. Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. J Pharmacol Exp Ther 1994;270(1):414–23.
[24] Yaeger R, Avila-Bront A, Abdul K, Nolan PC, Grann VR, Birchette MG, et al. Comparing genetic ancestry and self-described race in african americans born in the United States and in Africa. Cancer Epidemiol Biomarkers Prev 2008;17(6):1329–38.
[25] Race, Ethnicity, and Genetics Working Group. The use of racial, ethnic, and ancestral categories in human genetics research. Am J Hum Genet 2005;77(4):519–32.
[26] Winkler CA, Nelson GW, Smith MW. Admixture mapping comes of age. Annu Rev Genomics Hum Genet 2010;11:65–89.
[27] Thomas JH. Rapid birth-death evolution specific to xenobiotic cytochrome P450 genes in vertebrates. PLoS Genet 2007;3(5):e67.
[28] Sistkova J, Hudecek J, Hodek P, Frei E, Schmeiser HH, Stiborova M. Human cytochromes P450 1A1 and 1A2 participate in detoxication of carcinogenic aristolochic acid. Neuro Endocrinol Lett 2008;29(5):733–7.
[29] Wojnowski L, Turner PC, Pedersen B, Hustert E, Brockmöller J, Mendy M, et al. Increased levels of aflatoxin-albumin adducts are associated with CYP3A5 polymorphisms in The Gambia, West Africa. Pharmacogenetics 2004;14(10):691–700.
[30] Pedersen RS, Brasch-Andersen C, Sim SC, Bergmann TK, Halling J, Petersen MS, et al. Linkage disequilibrium between the CYP2C19*17 allele and wildtype CYP2C8 and CYP2C9 alleles: identification of CYP2C haplotypes in healthy Nordic populations. Eur J Clin Pharmacol 2010;66(12):1199–205.
[31] Carrillo JA, Christensen M, Ramos SI, Alm C, Dahl ML, Benitez J, et al. Evaluation of caffeine as an *in vivo* probe for CYP1A2 using measurements in plasma, saliva, and urine. Ther Drug Monit 2000;22(4):409–17.
[32] FDA U. Clinical pharmacogenomics: premarket evaluation in early-phase clinical studies and recommendations for labeling. US FDA, <www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM337169.pdf>; 2013 [cited 7.01.13].
[33] Frank D, Jaehde U, Fuhr U. Evaluation of probe drugs and pharmacokinetic metrics for CYP2D6 phenotyping. Eur J Clin Pharmacol 2007;63(4):321–33.
[34] Fuhr U, Jetter A, Kirchheiner J. Appropriate phenotyping procedures for drug metabolizing enzymes and transporters in humans and their simultaneous use in the "cocktail" approach. Clin Pharmacol Ther 2007;81(2):270–83.
[35] Bauer T, Bouman HJ, van Werkum JW, Ford NF, ten Berg JM, Taubert D. Impact of CYP2C19 variant genotypes on clinical efficacy of antiplatelet treatment with clopidogrel: systematic review and meta-analysis. BMJ 2011;343:d4588.
[36] Scott SA, Sangkuhl K, Stein CM, Hulot JS, Mega JL, Roden DM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. Clin Pharmacol Ther 2013;94(3):317–23.
[37] Wallentin L, James S, Storey RF, Armstrong M, Barratt BJ, Horrow J, et al. Effect of CYP2C19 and ABCB1 single nucleotide polymorphisms on outcomes of treatment with ticagrelor versus clopidogrel for acute coronary syndromes: a genetic substudy of the PLATO trial. Lancet 2010;376(9749):1320–8.
[38] Lee SS, Lee SJ, Gwak J, Jung HJ, Thi-Le H, Song IS, et al. Comparisons of CYP2C19 genetic polymorphisms between Korean and Vietnamese populations. Ther Drug Monit 2007;29(4):455–9.
[39] Yi B, Yang JY, Yang M. Past and future applications of CYP450-genetic polymorphisms for biomonitoring of environmental toxicants. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev 2007;25(4):353–77.
[40] Zhou SF, Yang LP, Zhou ZW, Liu YH, Chan E. Insights into the substrate specificity, inhibitors, regulation, and polymorphisms and the clinical impact of human cytochrome P450 1A2. AAPS J 2009;11(3):481–94.
[41] Hisaka A, Ohno Y, Yamamoto T, Suzuki H. Prediction of pharmacokinetic drug-drug interaction caused by changes in cytochrome P450 activity using *in vivo* information. Pharmacol Ther 2010;125(2):230–48.
[42] Ghotbi R, Gomez A, Milani L, Tybring G, Syvänen AC, Bertilsson L, et al. Allele-specific expression and gene methylation in the control of CYP1A2 mRNA level in human livers. Pharmacogenomics J 2009;9(3):208–17.
[43] Gunes A, Dahl ML. Variation in CYP1A2 activity and its clinical implications: influence of environmental factors and genetic polymorphisms. Pharmacogenomics 2008;9(5):625–37.
[44] Klein K, Winter S, Turpeinen M, Schwab M, Zanger UM. Pathway-targeted pharmacogenomics of CYP1A2 in human liver. Front Pharmacol 2010;1:129.
[45] Jurica J, Sulcova A. Topics on drug metabolism. InTech. <www.intechopen.com/books/topics-on-drug-metabolism/determination-of-cytochrome-p450-metabolic-activity-using-selective-markers>; 2012 [cited 2013].
[46] Ghotbi R, Christensen M, Roh HK, Ingelman-Sundberg M, Aklillu E, Bertilsson L. Comparisons of CYP1A2 genetic polymorphisms, enzyme activity and the genotype-phenotype relationship in Swedes and Koreans. Eur J Clin Pharmacol 2007;63(6):537–46.
[47] Han XM, Ouyang DS, Chen XP, Shu Y, Jiang CH, Tan ZR, et al. Inducibility of CYP1A2 by omeprazole *in vivo* related to the genetic polymorphism of CYP1A2. Br J Clin Pharmacol 2002;54(5):540–3.
[48] Sachse C, Brockmoller J, Bauer S, Roots I. Functional significance of a C–>A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine. Br J Clin Pharmacol 1999;47(4):445–9.
[49] Aklillu E, Carrillo JA, Makonnen E, Hellman K, Pitarque M, Bertilsson L, et al. Genetic polymorphism of CYP1A2 in Ethiopians affecting induction and expression: characterization of novel haplotypes with single-nucleotide polymorphisms in intron 1. Mol Pharmacol 2003;64(3):659–69.
[50] Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. Clin Pharmacol Ther 2012;92(4):414–7.
[51] Hisaka A, Ohno Y, Yamamoto T, Suzuki H. Theoretical considerations on quantitative prediction of drug-drug interactions. Drug Metab Pharmacokinet 2010;25(1):48–61.

[52] Garcia-Martin E, Martinez C, Ladero JM, Agundez JA. Interethnic and intraethnic variability of CYP2C8 and CYP2C9 polymorphisms in healthy individuals. Mol Diagn Ther 2006;10(1):29–40.

[53] Rodríguez-Antona C, Niemi M, Backman JT, Kajosaari LI, Neuvonen PJ, Robledo M, et al. Characterization of novel CYP2C8 haplotypes and their contribution to paclitaxel and repaglinide metabolism. Pharmacogenomics J 2008;8(4):268–77.

[54] Romkes M, Faletto MB, Blaisdell JA, Raucy JL, Goldstein JA. Cloning and expression of complementary DNAs for multiple members of the human cytochrome P450IIC subfamily. Biochemistry 1991;30(13):3247–55.

[55] Limdi N, Goldstein J, Blaisdell J, Beasley T, Rivers C, Acton R. Influence of CYP2C9 Genotype on warfarin dose among African American and European Americans. Per Med 2007;4(2):157–69.

[56] Scott SA, Jaremko M, Lubitz SA, Kornreich R, Halperin JL, Desnick RJ. CYP2C9*8 is prevalent among African-Americans: implications for pharmacogenetic dosing. Pharmacogenomics 2009;10(8):1243–55.

[57] Ingelman-Sundberg M, Daly AK, Nebert DW. Home page of the human cytochrome P450 (CYP) allele nomenclature committee, <www.cypalleles.ki.se/>; 2010 [cited 29.08.10].

[58] Xie HG, Kim RB, Wood AJ, Stein CM. Molecular basis of ethnic differences in drug disposition and response. Annu Rev Pharmacol Toxicol 2001;41:815–50.

[59] Cai WM, Nikoloff DM, Pan RM, de Leon J, Fanti P, Fairchild M, et al. CYP2D6 genetic variation in healthy adults and psychiatric African-American subjects: implications for clinical practice and genetic testing. Pharmacogenomics J 2006;6(5):343–50.

[60] Neafsey P, Ginsberg G, Hattis D, Sonawane B. Genetic polymorphism in cytochrome P450 2D6 (CYP2D6): population distribution of CYP2D6 activity. J Toxicol Environ Health B Crit Rev 2009;12(5–6):334–61.

[61] Sistonen J, Fuselli S, Palo JU, Chauhan N, Padh H, Sajantila A. Pharmacogenetic variation at CYP2C9, CYP2C19, and CYP2D6 at global and microgeographic scales. Pharmacogenet Genomics 2009;19(2):170–9.

[62] Lakhman SS, Ma Q, Morse GD. Pharmacogenomics of CYP3A: considerations for HIV treatment. Pharmacogenomics 2009;10(8):1323–39.

[63] Elens L, Bouamar R, Hesselink DA, Haufroid V, van der Heiden IP, van Gelder T, et al. A new functional CYP3A4 intron 6 polymorphism significantly affects tacrolimus pharmacokinetics in kidney transplant recipients. Clin Chem 2011;57(11):1574–83.

[64] Wang D, Guo Y, Wrighton SA, Cooke GE, Sadee W. Intronic polymorphism in CYP3A4 affects hepatic expression and response to statin drugs. Pharmacogenomics J 2011;11(4):274–86.

[65] Thompson EE, Kuttab-Boulos H, Witonsky D, Yang L, Roe BA, Di Rienzo A. CYP3A variation and the evolution of salt-sensitivity variants. Am J Hum Genet 2004;75(6):1059–69.

[66] Mirghani RA, Sayi J, Aklillu E, Allqvist A, Jande M, Wennerholm A, et al. CYP3A5 genotype has significant effect on quinine 3-hydroxylation in Tanzanians, who have lower total CYP3A activity than a Swedish population. Pharmacogenet Genomics 2006;16(9):637–45.

[67] Wilkinson GR. Genetic variability in cytochrome P450 3A5 and *in vivo* cytochrome P450 3A activity: some answers but still questions. Clin Pharmacol Ther 2004;76(2):99–103.

[68] Liu YT, Hao HP, Liu CX, Wang GJ, Xie HG. Drugs as CYP3A probes, inducers, and inhibitors. Drug Metab Rev 2007;39(4):699–721.

[69] Hart SN, Li Y, Nakamoto K, Wesselman C, Zhong XB. Novel SNPs in cytochrome P450 oxidoreductase. Drug Metab Pharmacokinet 2007;22(4):322–6.

[70] Hart SN, Zhong XB. P450 oxidoreductase: genetic polymorphisms and implications for drug metabolism and toxicity. Expert Opin Drug Metab Toxicol 2008;4(4):439–52.

[71] Ogasawara K, Chitnis SD, Gohh RY, Christians U, Akhlaghi F. Multidrug resistance-associated protein 2 (MRP2/ABCC2) haplotypes significantly affect the pharmacokinetics of tacrolimus in kidney transplant recipients. Clin Pharmacokinet 2013;52(9):751–62.

[72] Drögemöller B, Plummer M, Korkie L, Agenbag G, Dunaiski A, Niehaus D, et al. Characterization of the genetic variation present in CYP3A4 in three South African populations. Front Genet 2013;4:17.

[73] Xie HG, Wood AJ, Kim RB, Stein CM, Wilkinson GR. Genetic variability in CYP3A5 and its possible consequences. Pharmacogenomics 2004;5(3):243–72.

[74] Diczfalusy U, Miura J, Roh HK, Mirghani RA, Sayi J, Larsson H, et al. 4Beta-hydroxycholesterol is a new endogenous CYP3A marker: relationship to CYP3A5 genotype, quinine 3-hydroxylation and sex in Koreans, Swedes and Tanzanians. Pharmacogenet Genomics 2008;18(3):201–8.

[75] Pavek P, Dvorak Z. Xenobiotic-induced transcriptional regulation of xenobiotic metabolizing enzymes of the cytochrome P450 superfamily in human extrahepatic tissues. Curr Drug Metab 2008;9(2):129–43.

[76] Di YM, Chow VD-W, Yang L-P, Zhou S-F. Structure, function, regulation and polymorphism of human cytochrome P450 2A6. Curr Drug Metab 2009;10(7):754–80.

[77] Oscarson M, McLellan RA, Gullsten H, Agundez JA, Benitez J, Rautio A, et al. Identification and characterisation of novel polymorphisms in the CYP2A locus: implications for nicotine metabolism. FEBS Lett 1999;460(2):321–7.

[78] Tricker AR. Nicotine metabolism, human drug metabolism polymorphisms, and smoking behaviour. Toxicology 2003;183(1–3):151–73.

[79] Mwenifumbo JC, Zhou Q, Benowitz NL, Sellers EM, Tyndale RF. New CYP2A6 gene deletion and conversion variants in a population of Black African descent. Pharmacogenomics 2010;11(2):189–98.

[80] Mwenifumbo JC, Al Koudsi N, Ho MK, Zhou Q, Hoffmann EB, Sellers EM, et al. Novel and established CYP2A6 alleles impair *in vivo* nicotine metabolism in a population of Black African descent. Hum Mutat 2008;29(5):679–88.

[81] Moolchan ET, Franken FH, Jaszyna-Gasior M. Adolescent nicotine metabolism: ethnoracial differences among dependent smokers. Ethn Dis 2006;16(1):239–43.

[82] Nakajima M, Kuroiwa Y, Yokoi T. Interindividual differences in nicotine metabolism and genetic polymorphisms of human CYP2A6. Drug Metab Rev 2002;34(4):865–77.

[83] Schoedel KA, Hoffmann EB, Rao Y, Sellers EM, Tyndale RF. Ethnic variation in CYP2A6 and association of genetically slow nicotine metabolism and smoking in adult Caucasians. Pharmacogenetics 2004;14(9):615–26.

[84] McGraw Sr JE, Waller DP. Specific human CYP 450 isoform metabolism of a pentachlorobiphenyl (PCB-IUPAC# 101). Biochem Biophys Res Commun 2006;344(1):129–33.

[85] McGraw JE, Waller DP. The role of african american ethnicity and metabolism in sentinel polychlorinated biphenyl congener serum levels. Environ Toxicol Pharmacol 2009;27(1):54–61.

[86] Wang H, Tompkins LM. CYP2B6: new insights into a historically overlooked cytochrome P450 isozyme. Curr Drug Metab 2008;9(7):598–610.

[87] Zanger UM, Klein K. Pharmacogenetics of cytochrome P450 2B6 (CYP2B6): advances on polymorphisms, mechanisms, and clinical relevance. Front Genet 2013;4:24.

[88] Ilic K, Hawke RL, Thirumaran RK, Schuetz EG, Hull JH, Kashuba AD, et al. The influence of sex, ethnicity, and CYP2B6 genotype on bupropion metabolism as an index of hepatic CYP2B6 activity in humans. Drug Metab Dispos 2013;41(3):575–81.

[89] Ring HZ, Valdes AM, Nishita DM, Prasad S, Jacob 3rd P, Tyndale RF, et al. Gene-gene interactions between CYP2B6 and CYP2A6 in nicotine metabolism. Pharmacogenet Genomics 2007;17(12):1007–15.

[90] Hofmann MH, Blievernicht JK, Klein K, Saussele T, Schaeffeler E, Schwab M, et al. Aberrant splicing caused by single nucleotide polymorphism c.516G>T [Q172H], a marker of CYP2B6*6, is responsible for decreased expression and activity of CYP2B6 in liver. J Pharmacol Exp Ther 2008;325(1):284–92.

[91] McGraw J, Waller D. Cytochrome P450 variations in different ethnic populations. Expert Opin Drug Metab Toxicol 2012;8(3):371–82.

[92] Rohrbacher M, Kirchhof A, Geisslinger G, Lotsch J. Pyrosequencing-based screening for genetic polymorphisms in cytochrome P450 2B6 of potential clinical relevance. Pharmacogenomics 2006;7(7):995–1002.

[93] Howard LA, Ahluwalia JS, Lin SK, Sellers EM, Tyndale RF. CYP2E1*1D regulatory polymorphism: association with alcohol and nicotine dependence. Pharmacogenetics 2003;13(6):321–8.

[94] Stingl JC, Brockmöller J, Viviani R. Genetic variability of drug-metabolizing enzymes: the dual impact on psychiatric therapy and regulation of brain function. Mol Psychiatry 2013;18(3):273–87.

[95] Haufroid V, Jakubowski M, Janasik B, Ligocka D, Buchet JP, Bergamaschi E, et al. Interest of genotyping and phenotyping of drug-metabolizing enzymes for the interpretation of biological monitoring of exposure to styrene. Pharmacogenetics 2002;12(9):691–702.

[96] Ganetsky M, Böhlke M, Pereira L, Williams D, LeDuc B, Guatam S, et al. Effect of excipients on acetaminophen metabolism and its implications for prevention of liver injury. J Clin Pharmacol 2013;53(4):413–20.

[97] McCarver DG, Byun R, Hines RN, Hichme M, Wegenek W. A genetic polymorphism in the regulatory sequences of human CYP2E1: association with increased chlorzoxazone hydroxylation in the presence of obesity and ethanol intake. Toxicol Appl Pharmacol 1998;152(1):276–81.

[98] Lucas D, Ménez C, Girre C, Berthou F, Bodénez P, Joannet I, et al. Cytochrome P450 2E1 genotype and chlorzoxazone metabolism in healthy and alcoholic Caucasian subjects. Pharmacogenetics 1995;5(5):298–304.

[99] Bansal S, Anandatheerthavarada HK, Prabu GK, Milne GL, Martin MV, Guengerich FP, et al. Human cytochrome P450 2E1 mutations that alter mitochondrial targeting efficiency and susceptibility to ethanol-induced toxicity in cellular models. J Biol Chem 2013;288(18):12627–44.

Chapter 17

Pharmacokinetic Pharmacogenomics

Safaa Mohammed M. Alsanosi, Craig Skiffington, and Sandosh Padmanabhan
BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences,
University of Glasgow, UK

Chapter Outline

17.1 Overview 341
17.2 Principles of Pharmacokinetics 341
17.2.1 Absorption 341
17.2.2 Distribution 342
17.2.3 Metabolism 343
17.2.4 Excretion 346
17.2.5 Half-Life 348
17.2.6 Clearance 349
17.2.7 Zero- and First-Order Kinetics 349
17.3 ADME: Pharmacogenomics 349
17.3.1 CYP450 Enzymes 350
17.3.2 Non-CYP450 Enzymes 354
17.3.3 Drug Transporters 357
17.4 Conclusions 360
References 361

17.1 OVERVIEW

The goal of therapeutics is to achieve a definite beneficial effect with minimal adverse effect, and the approach to this has been through combining the principles of pharmacokinetics and pharmacodynamics. Pharmacokinetics (PK) is the branch of pharmacology that deals with the absorption, distribution, and elimination of drugs, while pharmacodynamics (PD) deals with the actions of drugs on the organism. PK governs the relationship between the drug dose and its concentration whereas PD relates the concentration to the effect; together PK/PD clarify a drug's dose–effect relationship. Since the 1950s there have been major advances in understanding the genetic basis of the interindividual variability observed in the pharmacokinetics, efficacy, and toxicity of various drugs. Recently, the European Medicines Agency (EMA) published guidance on the use of pharmacogenetics to investigate PK properties of new medicines [1] and PK pharmacogenetics feature in an increasing number of drug labels approved by the FDA [2]. Many pharmaceutical companies investigate the associations between PK gene variants and the observed interindividual pharmacokinetic and pharmacodynamic variability of drugs in early clinical development (phases I and II) [3]. In this chapter, we outline the basic principles of pharmacokinetics and summarize the progress in the pharmacogenomics of key pharmacokinetic systems.

17.2 PRINCIPLES OF PHARMACOKINETICS

PK studies are usually based on measuring the drug levels in blood and urine at different times following drug administration in order to understand how rapidly and for how long the drug appears in the target organ [4]. This involves mathematical modeling of data obtained from these studies around four domains: drug absorption (A), distribution (D), metabolism (M), and excretion (E). Together these are referred as ADME [4,5].

17.2.1 Absorption

Drugs are normally administered through various routes: orally, subcutaneously, intramuscularly, intravenously, rectally, or sublingually. Once a drug is administered, it is absorbed and distributed to different cells and organs. Drug absorption is generally defined as the rate and extent to

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00017-7
© 2014 Elsevier Inc. All rights reserved.

which the drug moves from its site of administration to its intended target (site) of action [6,7]. Absorption is a critical component in drug PKs, and several barriers must be overcome for a drug to reach its effect site. The rate and extent of absorption, as well as the time required to notice an observable effect, dictate the dose that should be administered [8].

Bioavailability indicates the proportion of the drug that passes into systemic circulation after administration, taking into account both absorption and local metabolic degradation [7]. Intravenous delivery provides 100% bioavailability, while an orally ingested drug may be incompletely absorbed by the GI tract before it reaches systemic circulation (around 75% of an orally administered drug is absorbed in 1–3 hours). In addition, the gastrointestinal lining expresses several cytochrome P450 enzymes and drug-efflux transporters (e.g., ATP-binding cassette, or ABC) that can decrease bioavailability [6,9].

Bioavailability depends on several factors, including the drug's physicochemical properties and formulation, first-pass metabolism, concomitant drug therapies, compliance, and disease state. The area under the curve (AUC) calculated from time zero to infinity following drug administration is an important PK parameter that measures the patient's "exposure" to a drug and depends on dose, bioavailability, and clearance [6]. The patient's plasma drug concentration time profile can be developed by measuring the plasma drug concentration at many time points and then estimating AUC, which is proportional to the absorbed fraction only when clearance is constant and the concentration is standardized.

Following linear kinetics, AUC is directly proportional to drug dose and inversely proportional to drug clearance. Therefore, the higher the clearance, the less time the drug spends in the systemic circulation and the earlier the decline in the concentration. As a result, the body's exposure to the drug and the AUC are smaller. Clearance is slightly dependent on the shape of the concentration time profile and is calculated by dividing the dose absorbed by the AUC [10].

Orally administered drugs are absorbed from the GI system and transferred to the liver via the portal vein before entering systemic circulation. The drug-metabolizing systems in the liver can thus exert a substantial effect on bioavailability; this is called first-pass metabolism or presystemic elimination. The larger the first-pass metabolism, the smaller the bioavailability of an orally administered drug. The liver drug metabolizing enzymes can be completely avoided by sublingual or buccal administrations or partially avoided by rectal administration [6]. Drugs that undergo extensive first-pass metabolism and cannot be administered orally (e.g., nitroglycerin) require alternative routes of administration, such as sublingual or intravenous. Drugs with extensive first-pass metabolisms can still be administered orally with higher doses. For example, a typical IV dose of verapamil is 1–5 mg, compared to the usual single oral dose of 40–120 mg.

First-pass metabolism is a major reason for recognizable differences in drug bioavailability among individuals, as even healthy people show considerable metabolizing differences in liver capacity. Moreover, in patients with severe liver disease, first-pass metabolism may be sharply decreased, leading to greater bioavailability [6]. Marked interindividual variations in terms of first-pass metabolism lead to unpredictable consequences. Examples of drugs that undergo significant "first-pass effect" are aspirin, morphine, levodopa, verapamil, salbutamol, and lidocaine [7,11].

17.2.2 Distribution

Once a drug has reached the bloodstream, it is distributed into the interstitial and intracellular compartments. As a general rule, the movement of lipophilic drugs is faster than that of hydrophilic drugs, and small lipophilic molecules distribute across cell membranes more easily than do large polar molecules [12–14]. Lipophobic drugs are primarily confined to plasma and interstitial fluids and generally do not go through the brain tissue after acute dosing. Hydrophilic drugs, such as the aminoglycosides, are mostly distributed into extracellular fluid, and their volume is affected by fluid retention or dehydration, both of which can happen in a number of renal diseases.

Passive diffusion, facilitated diffusion, and active transport are the three mechanisms involved in drug distribution across cell membranes [15]. Solute carrier (SLC) and ATP-binding cassette (ABC) transporters play an important role in active drug transport [16]. Cardiac output, regional blood flow, capillary permeability, lipid solubility, and plasma-protein binding determine the rate and amount of drug distributed into tissue [12].

Sites with high blood flow primarily receive greater amounts of a drug compared to sites with low or disturbed blood flow. Consequently, drug concentrations increase faster in organs such as the brain, heart, and kidneys as compared with skin, muscle, and bone. The structure and permeability of capillaries varies depending on the organ, and this affects how the drug is distributed. For example, capillaries in the kidney and liver sinusoids show more permeability as compared with the tight junctions between endothelial cells that line the brain capillaries, creating a relatively impermeable blood–brain barrier (BBB). The BBB allows only selective transport of lipophilic molecules and prevents passive entry of lipophobic/ionized molecules into the brain [12,13].

17.2.2.1 Plasma-Protein Binding

In blood, many drugs are bound to plasma protein. This binding is reversible with bound and unbound drug fractions in dynamic equilibrium. Any change in unbound-drug concentration is directly followed by a change in bound-drug concentration. Only free (unbound) fractions are able to cross membranes or interact with drug targets and are considered pharmacologically active. In general, the amount of a drug that is bound to protein depends on three main factors: free-drug concentration, binding site affinity, and protein concentration [17,18].

The most significant plasma-protein binding is to albumin, which is the major carrier of acidic drugs (e.g., warfarin, sulfonamides). Other plasma proteins, such as α1-acid glycoprotein, bind basic drugs. In addition, certain drugs may bind to proteins that function as specific hormone carrier proteins, such as the binding of estrogen or testosterone to sex hormone–binding globulin or the binding of thyroid hormone to thyroxin-binding globulin. For many drugs, more than 90% in the plasma is bound to a protein (e.g., warfarin, diazepam), whereas other drugs may have less extensive protein binding (digoxin and gentamicin). Because binding of drugs to plasma proteins such as albumin is nonselective, and because the number of binding sites is relatively large (high-capacity), many drugs with similar physicochemical characteristics can compete with each other and with endogenous substances for these binding sites, resulting in displacement of one drug by another and leading to an increase in its pharmacological activity or toxicity.

Changes in plasma-protein concentrations can also lead to variability in pharmacological effects of drugs, which are also affected by disease-related factors; for example, hypoalbuminemia secondary to nephrotic syndrome results in decreased drug binding, which increases the drug's unbound fraction of the drug. In addition, acute-phase reaction responses (e.g., cancer, arthritis) can lead to high levels of α1-acid glycoprotein and can enhance basic drug binding. Misinterpretation of plasma drug concentration measurements is a common problem that results from drugs competing for plasma-protein binding sites, as most assays do not differentiate unbound from bound drugs.

A number of drugs concentrate in body tissues at higher levels than those in extracellular fluids and blood; for example, the concentration of an antimalarial agent (quinacrine) in the liver can be many thousand times higher than that in the blood if the drug is administered for long periods. Reversible tissue binding of drugs regularly happens with proteins and phospholipids. Fractions of a drug in the body may be bound in this way and can act as a reservoir that prolongs drug action in the same tissue or at a distant site. Lipophilic drugs can accumulate in adipose tissue [12]. Halothane can concentrate in fat during long operations, and its slow release can cause postoperative prolonged central nervous system (CNS) depression. Tissue binding and drug accumulation may lead to local toxicity, as in the case of aminoglycoside antibiotic (gentamicin) accumulation in the kidneys and vestibular system [12].

17.2.2.2 Volume of Distribution

Volume of distribution (Vd), represents the apparent volume into which the drug is distributed to provide the same concentration as it currently is in blood plasma. It is calculated by the amount of the drug in the body divided by the plasma concentration [19]. Thus, Vd reflects the extent to which the drug is present in extravascular tissues but not in plasma. Lipid solubility can affect Vd, as highly lipid-soluble drugs have good cell penetration, resulting in high Vd. Plasma-protein binding, particularly to albumin, reduces the Vd, while tissue binding increases it [17].

Vd can be used to determine the size of a loading dose in order to quickly reach the required therapeutic plasma concentration, assuming that successful therapy is directly linked to the plasma concentration and that there are no adverse effects if a quite large dose is rapidly administered. In addition, Vd is helpful in predicting the initial maximum concentration for an IV bolus and can be used to predict the effectiveness of dialysis in treating drug intoxication during an emergency [18].

17.2.3 Metabolism

Once drugs have been distributed throughout the body, they are metabolized into more polar, inactive metabolites in order to eliminate them from the body. The lipophilic characteristics of drugs that promote their passage through biological membranes and subsequent access to their site of action also serve to hinder their excretion from the body. Metabolism includes two main processes: one, the molecule is made more lipophobic in order to reduce the possibility of reabsorption in the renal tubules; two, the molecule is conjugated to reduce its effect and increase its excretion. While for some drugs the metabolites have the actions of the parent drug (e.g., diazepam and its metabolite, nordiazepam), for others, the metabolite may result in toxicity (e.g., paracetamol). Also, a number of drugs are "prodrugs" (i.e., inactive) but transform into an active drug in the body. Prodrugs are often designed to improve oral bioavailability in cases where the intended drug is poorly absorbed through the gastrointestinal tract; for example, enalapril is hydrolyzed to the active compound enalaprilat [12].

Metabolizing enzymes are distributed in many tissues in the human body, with the highest levels found in GI tract tissues (e.g., the liver and the small and large intestines). The liver is the major metabolizing organ for both drugs and endogenous chemicals (e.g., cholesterol, fatty acids, and proteins). The small intestine also plays a vital role in drug metabolism, as most orally administered drugs become inactivated metabolically in either the intestinal epithelium or the liver before they reach systemic circulation [20]. Drug metabolism reactions are classified as either phase I or phase II.

17.2.3.1 Phase I Reactions

Drugs metabolized by phase I reactions usually lose their pharmacological activity, although there are some examples of enhanced or altered pharmacological activity. If not eliminated rapidly through urine, phase I reaction products can react with endogenous compounds to form a highly water-soluble conjugate [12].

Phase I reactions "functionalize" the drug for phase II, as they introduce a functional group—such as –OH, –COOH, –SH, –O–, or NH—which increases the polarity of the drug molecule and provides a site for phase II reactions. Phase I metabolic reactions include oxidation, reduction, and hydrolysis [20] (see Table 17.1).

Oxidations are the most common type of reaction and are catalyzed by the cytochrome P450 (CYP) system, which is located in the smooth endoplasmic reticulum. Metabolizing enzymes responsible for oxidation are mainly CYPs, alcohol dehydrogenase (ADH), aldehyde dehydrogenase (ALDH), dihydropyrimidine dehydrogenase (DPD), monoamine oxidase (MAO),

TABLE 17.1 Phase I and Phase II Reactions, Enzymes, and Substrate Drugs

Reaction	Enzymes	Drugs
Phase I		
Oxidation	N-dealkylation	Fentanyl, Buprenorphine, Morphine, Imipramine, Diazepam Codeine, Erythromycin, Tamoxifen, Theophylline, Caffeine, Methadone
	O-dealkylation	Codeine, Dextromethorphan, Indometacin,
	Aliphatic hydroxylation	Ibuprofen, Tolbutamide, Phenobarbital, Meprobamate, Cyclosporine, Midazolam
	Aromatic hydroxylation	Phenytoin, Phenobarbital, Propranolol, Ethinyl Estradiol, Amphetamine, Warfarin
	N-oxidation	Dapsone, Chlorpheniramine, Meperidine
	S-oxidation	Cimetidine, Chlorpromazine, Thioridazine, Omeprazole
	Deamination	Diazepam, Amphetamine
Hydrolysis		Modafinil, Salvinorin A, Carbamazepine, Procaine, Aspirin, Clofibrate Meperidine, Enalapril, Cocaine, Lidocaine, Procainamide, Indomethacin, Suxamethonium
Reduction		Prednisone
Phase II		
Glucuronidation	UDP-glucuronosyl transferases (UGT)	Morphine, Acetaminophen, Oxazepam, Lorazepam, UDP-Glucuronic
Sulphation	Sulfotransferases (SULT)	Paracetamol, Steroids, Methyldopa
Methylation	Methyltransferases (MT)	L-Dopa, Methyldopa, Mercaptopurine, Captopril
Acetylation	N-acetyltransferases (NAT)	Sulphonamides, Isoniazid, Dapsone, Clonazepam
Amino acid conjugation	Bile acid-coenzyme A:amino acid N-acyltransferase (BAAT)	Glycine
Glutathione conjugation	Glutathione-S-transferases (GST)	Paracetamol, Adriamycin, Fosfomycin, Busulfan

and flavin-containing monooxygenase (FMO) [20,21]. Oxidation via phase I enzymes adds or exposes a functional group, allowing phase I metabolites to act as substrates for the phase II conjugating enzyme [20]. The oxidation of a drug by the CYP system requires a P450 enzyme, molecular oxygen, nicotinamide adenine dinucleotide phosphate (NADPH), and a flavoprotein (NADPH-P450 reductase).

Hydrolysis is not limited to the liver but also takes place in various tissues. The main metabolic enzymes involved in drug hydrolysis are epoxide hydrolases (EH), esterases, and amidases [21]. Inactive prodrugs are converted rapidly to active metabolites mostly through an ester or amide linkage hydrolysis; for example, enalapril, the angiotensin-converting enzyme inhibitor, is quite inactive until it is transformed by esterase into its diacid metabolite (enalaprilat) [12]. Examples of drugs that undergo hydrolysis include carbamazepine, aspirin, meperidine, lidocaine, procainamide, and indomethacin [20].

Reduction reactions are much less common than oxidation reactions. The main metabolic enzyme involved in reduction is NADPH-CYP reductase [21]. A few drugs undergo reduction reactions, and some reductive reactions are important; for example, warfarin is inactivated by conversion of a ketone to a hydroxyl group by the CYP2A6 enzyme.

CYPs, FMO, and EH, in addition to some phase II conjugating enzymes, particularly Uridine 5′-diphospho-glucuronosyltransferase (UDP-glucuronosyltransferase, or UGT), are located mainly in the endoplasmic reticulum, while phase II enzymes are mostly cytosolic [20].

17.2.3.2 Phase II Reactions

Whereas various phase I reactions inactivate a drug biologically, phase II enzymes enable drug elimination by significantly increasing water solubility, abolishing pharmacological activity, and increasing molecular weight. Phase II reactions are generally biosynthetic (conjugation) reactions, such as glucuronidation, sulphation, acetylation, methylation, and glutathione conjugation [20]. These highly polar conjugates in general are inactive and are eliminated rapidly through urine and feces.

Phase II enzymes comprise a number of conjugating enzyme superfamilies, including the glutathione-S-transferases (GST), UGT, sulfotransferases (SULT), *N*-acetyltransferases (NAT), and methyltransferases (MT). Phase II reactions as a whole depend on catalytic reactions for cofactors such as UDP-glucuronic acid (UDP-GA) for UGT and 3-phosphoadenosine-5-phosphosulfate (PAPS) for SULT, which respond to the substrate functional groups. These functional groups are mainly formed via phase I CYPs. In cholestasis, toxic bile acids accumulate in hepatic cells, leading to their damage and functional impairment. Glucuronidation, which is catalyzed by UGT enzymes, is considered an essential metabolic pathway for hepatic bile acids [20,22].

17.2.3.3 Non-CYP450 Enzymes

Non-CYP450 enzymes play a vital role in both the metabolism and the elimination of several drugs; for example, UGT, SULT, thiopurine S-methyltransferase (TPMT), DPD, NAT, and GST have been indentified for their clinical significance [23].

UGTs are membrane-bound enzymes found in the hepatic endoplasmic reticulum and various other extrahepatic tissues [21]. They involve a superfamily of vital proteins that catalyze the glucuronidation reaction on a wide range of structurally different endogenous and exogenous compounds. UGTs catalyze the transfer of the glucuronic acid group of uridine diphosphoglucuronic acid to the functional group (e.g., hydroxyl, carboxyl, amino, sulfur) of a specific substrate [24].

SULTs are localized in the cytosol and catalyze the transfer of the sulfonyl group from the cofactor PAPS to the nucleophilic sites of a range of substrates, including hormones and drugs. In humans, 11 SULT isoforms have been recognized and have been classified into four main groups based on evolutionary projections (SULT 1, 2, 3, and 4) [20,21]. They play a significant role in human homeostasis. For instance, SULT1B1, which is mostly expressed in the skin and brain, is responsible for cholesterol and thyroid hormone catalysis [20].

TPMT is recognized for its important role in the metabolism of thiopurine drugs (e.g., 6-mercaptopurine (6-MP), azathiopurine (AZA), and 6-thioguanine) and is catalyzed through the s-methylation of aromatic and heterocyclic sulfhydryl [21]. AZA and 6-MP are used to treat inflammatory bowel disease and autoimmune disorders such as systemic lupus erythematosus and rheumatoid arthritis. 6-MP is also used for the treatment of childhood acute lymphoblastic leukemia (ALL). 6-thioguanine is used to treat acute myeloid leukemia (AML), and since TPMT is responsible for the detoxification of 6-MP, any deficiency in it can result in severe toxicities in patients taking those drugs [20,25]. DPD is the rate-limiting enzyme in pyrimidine and 5-fluorouracil (5-FU) degradation. 5-FU is commonly prescribed to treat GI malignancies, and DPD deficiency can increase concentrations of bioavailable 5-FU anabolic products, leading to 5-FU–related toxicity [26].

GST catalyzes glutathione conjugation to a broad range of endogenous metabolites and drugs. Human GSTs are classified into three major families: cytosolic/nuclear, mitochondrial, and microsomal. They also have nonenzymatic functions, as they work as regulators of cell signaling and the post-translational modification pathway in response to stress, growth factors, DNA damage, cell proliferation, cell death, and other processes that eventually lead to tumor growth and drug resistance. Because of their functionalities, GSTs are seen as significant determinants of cancer susceptibility, therapeutic response, and prognosis [20,21,27].

NATs are responsible for the metabolism of both drugs and environmental compounds that contain an aromatic amine or hydrazine group. NATs catalyze the transfer of an acetyl group from acetyl coenzyme A to arylamines, arylhydroxylamines, and arylhydrazines [20,21].

17.2.3.4 *Enzyme Induction*

The pharmacological effect of a drug is dependent on its concentration at its site of action, which is partly dependent on its metabolism rate. Any changes in this enzyme activity can affect drug action. Enzyme induction can be defined as the increased synthesis (higher amount) or decreased degradation (increased activity) of enzymes that occurs as a result of the presence of an exogenous substance [20,28]. It is usually associated with a reduction in drug efficacy but may also change the toxicity of particular substances. CYP3A4 plays a role in the metabolism of about 50% of the drugs that are currently prescribed. Induction of CYP3A4 by rifampicin can increase oestrogen metabolism, thus reducing the effectiveness of birth control pills.

Certain drugs can increase the rate of synthesis of CYP450 enzymes; consequently, this enzyme induction can enhance the clearance of other drugs. Generally, such induction needs exposure to the inducing agent for more than a week before effects can be observed [28]. Clinical studies have shown that carbamazepine, phenytoin, and rifampicin are the most potent enzyme inducers in clinical use as they produce many clinically significant drug interactions mainly associated with increases in the metabolism of CYP2C9, CYP2C19, and CYP3A4 substrates (see Table 17.2). However, enzyme induction is not limited to the administration of prescription drugs. St. John's Wort, an herbal medicine, can also induce metabolizing enzymes, as can tobacco, which induces a CYP1A2 substrate, such as theophylline [29].

17.2.3.5 *Enzyme Inhibition*

Enzyme inhibition refers to a decrease in enzyme-related processes, enzyme production, or enzyme activity. A number of clinically important interactions between drugs result from CYP450 inhibition. CYP450 inhibitors are different in their selectivity toward enzymes and are classified by their mechanisms of action. Some drugs are potent competitive inhibitors and compete for the active site, but they are not a substrate for the enzyme (e.g., quinidine and CYP2D6), while other drugs are noncompetitive inhibitors (e.g., ketoconazole and CYP3A4). Enzyme inhibition can cause many adverse drug interactions that tend to happen more rapidly (within a couple of days) than those seen with enzyme induction, as they occur once the concentration of the inhibiting drug becomes high enough to compete with the affected drug. Examples of enzyme-inhibiting agents are cimetidine, erythromycin, ciprofloxacin, and isoniazid.

In certain cases, enzyme inhibition can cause potentially serious adverse events; for example, ketoconazole reduces the metabolism of the CYP3A4 substrate (terfenadine), resulting in a prolonged QT interval and torsades de pointes. As with enzyme induction, enzyme inhibition is not limited to drug interactions [30]; for example, grapefruit juice (a CYP3A4 inhibitor) can cause clinically significant interactions with many drugs, such as midazolam, simvastatin, and terfenadine. Resulting in much higher plasma concentrations of the inhibited drug than intended, enzyme inhibition can be a major safety issue, such as in co-administration of ketoconazole or ritonavir with midazolam, which increases midazolam plasma exposure (AUC) by 15–20 times—a condition that should be avoided [8].

17.2.4 Excretion

Drugs are eliminated from the body through two main mechanisms—liver metabolism and renal excretion. Some drugs are excreted in insignificant amounts via other routes, such as sweat, saliva, and tears, and elimination through these routes depends mostly on nonionized lipophilic diffusion through the epithelial cells of the glands and on urine pH [12].

17.2.4.1 *Renal Excretion*

The kidney is the principal organ of excretion for a drug and its metabolites. Drugs that are water soluble are mostly excreted unchanged through the kidneys. Lipid-soluble drugs are not eliminated efficiently by the kidneys and first need to be metabolized to more polar products. Renal disease influences the excretion of particular drugs. The extent to which excretion is impaired can be deduced by measuring creatinine clearance. Drug and metabolite excretion in urine includes three different processes: glomerular filtration, active tubular secretion, and passive tubular reabsorption; changes in kidney function in general affect all three processes to a similar degree [12].

Glomerular filtration is the most common route of renal elimination. The amount of drug entering the tubular lumen through filtration depends on the glomerular filtration rate (GFR) and the level of plasma binding of the drug. Only unbound drugs are cleared by filtration; protein-bound drugs remain in circulation, where some of them dissociate to restore equilibrium. The size of the molecule is the only limiting factor at this step. In relation to filtered drugs, it is expected that great declines in renal function, as reflected by a decreased GFR, lead to drug accumulation, and therefore renal patients may need dose adjustments. However, for actively secreted drugs, GFR is less important than renal plasma flow [31].

TABLE 17.2 Cytochrome P450 Enzymes, Substrates, Inducers, and Inhibitors

Enzymes	Substrates	Inducers	Inhibitors
CYP2D6	Timolol, Metoprolol, Carvedilol, Fluoxetine, Paroxetine, Tricyclic antidepressants, Propafenone, Flecainide, Codeine, Phenformin, Codeine, Debrisoquine		Cimetidine, Quinidine, Halperidol, Bupropion, Fluoxetine, Paroxetine, Perhexaline, Cinacalcet, Doxepin, Duloxetine, Flecainide, Moclobemide, Quinine, Terbinafine, Sulphaphenazole, Benzoflavone, Troleandomycin, Ritonavir, Chlorpromazine, Miconazole, Diphenhydramine, Ketoconazole, Methadone, Nicardipine, Sertraline, Venlafaxine, Tricyclic antidepressants
CYP2C19	Omeprazole, Mephenytoin, Phenytoin	Rifampicin, Phenobarbital, Lopinavir/Ritonavir, St John's wort	Omeprazole, Fluconazole, Fluvoxamine, Ticlopidine, Clarithromycin, Fluoxetine, Moclobemide, Moriconazole, Diethyldithiocarbamate, Cimetidine, Ketoconazole
CYP1A2	Caffeine, Paracetamol, Tacrine, Theophylline	Phenytoin, Tobacco smoke, Phenobarbital, Beta-Naphthoflavone, Rifampin	Cimetidine, Ranitidine, Ciprofloxacin, Fluvoxamine, Ethinylestradiol, Interferon Alpha-2b, Methoxsalen, Sulphaphenazole, Benzoflavone, Troleandomycin, Furafylline
CYP2E1	Alcohol, Paracetamol	Isoniazid, Phenobarbital, Rifampin	Tetrahydrofurane, Diethyldithiocarbamate, Disulfiram, Sulphaphenazole, Benzoflavone, Quinidine, Troleandomycin
CYP2C8	Paclitaxel, Repaglinide, Rosiglitazone, Pioglitazone, Troglitazone, Amiodarone, Chloroquine, Amodiaquine, Verapamil, Ibuprofen, Ofluvastatin, Amitriptyline, Perphenazine, Diclofenac, Gallopamil, Omeprazole, Carbamazepine	Phenytoin, Rifampicin	Sulphaphenazole, Benzoflavone, Quinidine, Troleandomycin
CYP2C9	Warfarin, Phenytoin, Glipizide, Losartan, Ibuprofen, Tolbutamide	Rifampin Phenobarbital	Fluconazole, Sulphaphenazole, Benzoflavone, Quinidine, Troleandomycin, Ketoconazole, Amiodarone, Phenytoin
CYP3A	Calcium Channel Blockers, Antiarrhythmics (Lidocaine, Quinidine, Mexiletine), HMG-Coa Reductase Inhibitors, Cyclosporine, Tacrolimusindinavir, Saquinavir, Ritonavir, Nifedipine, Simvastatin	Bosentan, Rifampicin, Carbamazepine, Efavirenz, Phenobarbital, Modafinil, Phenytoin, St John's Wort, Clotrimazole, Ritonavir, Sulfinpyrazone	Erythromycin, Clarithromycin, Itraconazole, Ketoconazole, Lopinavir/Ritonavir, Ritonavir, Saquinavir, Saquinavir/Ritonavir, Voriconazole, Grapefruit Juice, Fluconazole, Amiodarone Cimetidine, Fluvoxamine, Diltiazem, Cyclosporin, Imatinib, Verapamil, Atazanavir/ Ritonavir, Atazanavir, Aprepitant, Sulphaphenazole, Benzoflavone, Quinidine, Troleandomycin, Diltiazem, Itraconazole
CYP2A6	Coumarin, Halothan, Methoxyflurane, Valporic Acid, Disulfiram, Losigamone, Letrozole	Phenobarbital, Rifampin	Methoxsalen, Diethyldithiocarbamate
CYP2B6	Cyclophosphamide, Methadone	Phenobarbital, Rifampin, Clotrimazole, Ritonavir, Sulfinpyrazone	Ritonavir, Efavirenz, Nelfinavir

Passive reabsorption in the distal tubule occurs only with unionized (lipid-soluble) drugs. Urine PH determines whether weak acids and bases are reabsorbed, and this in turn determines the degree of ionization. If renal function is impaired, for instance, by disease, the clearance of drugs that usually undergo renal excretion is decreased. Tubular reabsorption follows the role of passive nonionic diffusion and depends on both urinary pH and the drug's pKa. A pH that prefers the nonionized state of a drug increases both its lipophilicity and passive reabsorption across the tubular membrane. The extent of passive reabsorption also depends on urinary volume. The clinical relevance of renal tubular reabsorption relates generally to the disease effect on the process [31,32].

Active secretion in the proximal tubule is predominant in that the proximal convoluted tubule can actively secrete some compounds—mostly weak acids and bases—into the lumen of the nephron. Actively secreted drugs accumulate in patients with reduced renal function, as renal blood flow and GFR generally also decrease. The SLC transporters, which include organic cation transporters (OCTs) and organic anion transporters (OATs), are involved in active transport of many endogenous substances and drugs [32]. Actively secreted drugs are subject to competition for transport, so there are some concerns with particular clinical conditions in which co-administered drugs inhibit transport, which can expose patients to high drug concentrations [31]. As about 80% of a drug delivered to the kidney is subjected to the carrier, tubular secretion is probably the most efficient mechanism of renal drug elimination. A number of organic acids are actively secreted (e.g., penicillins, probenecid), as well as quite a few diuretics (e.g., furosemide, thiazides) and different conjugates (e.g., glucuronic acid, sulphate). Some organic cations are also secreted, such as morphine and amiloride.

17.2.4.2 Biliary Excretion

Liver cells transfer different substances, and bile excretion is a main route of excretion for a small number of drugs (e.g., cromoglycate) and a larger number of drug metabolites (e.g., morphine glucuronide). Drug conjugates are excreted into the bile and later released into the intestines, where they are reabsorbed into the body (enterohepatic circulation). Enterohepatic circulation has important clinical consequences, as it prolongs the drug effect and increases its plasma half-life ($t_{1/2}$). While some drugs are excreted mostly unchanged in bile (e.g., vecuronium), other drugs are excreted in bile after deacetylation, which retains their biological activity and stops their reabsorption. Drugs pass from plasma to bile via transport systems similar to renal tubules, such as OCTs, OATs, and P-glycoproteins.

17.2.5 Half-Life

The half-life ($t_{1/2}$) is one of the simplest PK parameters to calculate. It is the time taken for a drug concentration to fall to half of its original value, measured in hours. However, not only does it determine the time needed for drug concentrations to fall to immeasurably low levels following a single bolus; it is also the main determinant of the time required to achieve steady-state plasma concentrations (Css) after any changes in dose. When a drug is administered chronically and its amount administered per unit time equals the amount eliminated per unit time, this situation represents a steady state. While Css is stable in a continuous IV infusion, it varies during the dosing interval in chronic oral administration although the profile of time-concentration between dosing intervals is still stable.

The $t_{1/2}$ is related to elimination rate constant, clearance, and Vd. The "initial" concentration achieved after a single IV bolus, then, decreases by a constant amount per unit time as the drug is eliminated from the body. The parameter that describes this rate of decline is the elimination rate constant (k), and it depends on both clearance and volume of distribution. Clearance measures the ability of the body to eliminate a drug; therefore, as clearance decreases (e.g., with disease), $t_{1/2}$ is expected to increase. However, this kind of relationship is applicable only when the disease does not affect the Vd; for example, diazepam $t_{1/2}$ increases with patient age, but it is the volume of distribution that changes, not the clearance [12,18]. The $t_{1/2}$ represents the time taken to eliminate a drug and the time taken for the drug to accumulate to a steady state with multiple dosing or during a constant-rate infusion. This takes about 4–5 $t_{1/2}$ (when starting from zero). Drug concentration data are generally plotted against time. A drug achieves 50% of its Css after one $t_{1/2}$, 75% after two $t_{1/2}$, 88% after three $t_{1/2}$, 94% after four $t_{1/2}$, and 97% after five $t_{1/2}$.

$t_{1/2}$ is an important determinant of dose frequency and time required to achieve Css; however, it provides little information about differences in dose requirements associated with clinical conditions or diseases. For instance, a change in $t_{1/2}$ might reflect a change in Vd or clearance or both. In the same way, if both clearance and volume change in proportion, the $t_{1/2}$ may be unaltered although average steady-state concentrations may change. Clearance and volume should therefore be calculated individually whenever possible [18,33].

In a number of diseases, clearance and distribution volume can be affected by changes in protein binding, resulting in unpredictable $t_{1/2}$ changes. For example, in patients with acute viral hepatitis, tolbutamide $t_{1/2}$ exhibits enhanced clearance which is unexpected. This occurs because the disease changes both plasma and tissues protein binding, does not affect volume of distribution, but increases clearance because free drugs are present in the bloodstream in higher concentrations. Although it can be a poor indicator of drug elimination from the body, $t_{1/2}$ does give a good indication of the time required to reach a steady state whenever a drug regimen is initiated or changed, the time needed for a drug to be cleared from the body, and a means to predict the appropriate dosing interval [12,17].

17.2.6 Clearance

Clearance (CL) is the most important concept to consider when designing a rational regimen for long-term drug administration. CL (expressed as volume/time) describes the volume of fluid that is completely cleared of drug per unit time, mainly through hepatic metabolism and renal excretion [12]. While elimination $t_{1/2}$ determines the time required to achieve Css, the level of that steady state is determined through CL and dose alone. If the drug is given for long enough, the amount of drug eliminated from the body during one dosage interval is equal to the amount of drug that enters the systemic circulation during each dosage interval. This is known as a steady state. CL depends essentially on liver and/or kidney efficiency to eliminate a drug, and it varies in certain diseases that affect those organs or the blood supply to them. In a stable clinical situation, CL remains constant and is directly proportional to dose rate. The clinical importance of this is that, for most drugs, if the dose rate is doubled, the $Css_{average}$ doubles, and, if the dose rate is halved, the $Css_{average}$ is halved.

The concept of CL is very useful in clinical PKs because its value for certain drugs is generally constant over the clinically encountered concentration ranges. Drug elimination systems, such as metabolizing enzymes and transporters, are usually not saturated; therefore, the absolute elimination rate is basically a linear function of the drug's concentration in plasma. To be precise, the elimination of most drugs follows first-order kinetics, where a constant fraction of drug in the body is eliminated per unit of time. If the elimination mechanisms of a particular drug become saturated, the kinetics move toward zero-order kinetics, in which a constant amount of drug is eliminated per unit of time. In such situations, CL changes with drug concentration [12].

It is vital to understand that CL does not indicate how much of a drug is being cleared; rather, it indicates the volume of blood or plasma from which the drug would have to be totally removed to account for the CL. The CL by different organs (GI tract, kidney, liver, and other organs) is additive. Diseases, drug interactions, or even genetic variants that decrease the activity of drug-metabolizing enzymes or mechanisms of execration may decrease CL; consequently, a dose adjustment to avoid drug toxicity is required. On the other hand, a number of drug interactions and genetic variants increase CYP expression, and therefore increased drug dosage might be required to maintain a therapeutic effect [12].

17.2.7 Zero- and First-Order Kinetics

The hallmark of first-order (linear) kinetics is the proportionality between dose rate and Css. Most drugs show first-order kinetics, where a constant fraction of drug in the body is eliminated per unit time. Linear kinetics explains that the decrease in drug levels in the body is dependent on the plasma concentration (a concentration-dependent process). The higher the drug concentration, the larger the absolute amount of drug eliminated per unit time. Therefore, the rate of elimination is proportional to the amount of drug in the body, while CL remains constant. Generally, for drugs with first-order kinetics, Vd, $t_{1/2}$, k, and CL are all interrelated [12].

However, if the elimination mechanisms of a particular drug become saturated, the kinetics move toward zero-order (nonlinear) kinetics, in which a constant amount of drug is eliminated per unit of time. Various drugs follow zero-order kinetics at high or toxic concentrations. Metabolism in general, which involves particular enzymes, is one of the most important elements contributing to a drug undergoing zero-order kinetics. When metabolism enzymes reach a point of saturation, the rate of elimination does not increase in response to a concentration increase but rather becomes constant (a concentration-dependent process).

Common examples of drugs that follow nonlinear kinetics are aspirin, phenytoin, alcohol, heparin, and ethanol. The $t_{1/2}$ is not constant for zero-order reactions, but rather depends on concentration. The higher the concentration, the longer the $t_{1/2}$. The clinical significance of nonlinear kinetics is that any small increase in dose can cause a large increase in concentration. This is mainly important when toxic side effects and concentration are strongly related, which is the situation with, for example, phenytoin [12].

17.3 ADME: PHARMACOGENOMICS

As discussed previously, there are many sources of variation in enzyme activity; age, enzyme induction or inhibition, and diseases (especially of the liver) are among them. Variation in the DNA sequence of genes encoding enzymes can abolish, reduce, or increase the expression and activity of an enzyme, and this can manifest as the "metabolizer" phenotype in an individual. Individuals who are homozygous for the two alleles coding for "normal" enzyme function are termed extensive metabolizers (homozygous EM or "wild-type"); those who are homozygous with two variant alleles resulting in inactive or absent enzymes are "poor metabolizers" (PM); those who are heterozygous manifest an intermediate metabolizer (IM) phenotype with reduced function (heterozygous EM). Intermediate and extensive metabolizers are often collectively referred as extensive metabolizers, especially in studies in which metabolizer status is assigned using phenotype.

Gene duplication or multiplication, as, for example, seen in *CYP2D6*, can result in "ultrarapid metabolizer" (UM) phenotype. Standard drug doses achieve normal concentrations and effect in homozygous EMs (which usually make up the largest proportion of the population), but they may be toxic in PMs (possibly in heterozygous EMs or IMs) and ineffective in UMs, who may require a higher dose to achieve therapeutic effective drug concentrations [9,21,34].

In general, the most important factor affecting drug action is the AUC at the site of action, reflecting the "metabolizer" phenotype; AUC is the best pharmacokinetic end point for the assessment of pharmacogenetic effects. However, even if there is a clear association between concentration and drug effect/toxicity, there are multiple factors that determine the clinical relevance of a functional polymorphism in a drug metabolizing gene [34]. These are summarized here:

- Clear association between concentration and drug effect/toxicity. Genetic variation that affects the functioning of the drug metabolizing gene can affect the AUC and hence can predict drug effects.
- The clinical significance of a genetic variant depends on the drug's therapeutic index (the ratio between dose effectiveness and safety). A high therapeutic index indicating that a drug is safe over a wide range of concentrations and variations due to genetics may be irrelevant, whereas a low therapeutic index indicates that minor variations in drug concentration, such as from polymorphisms, may be important.
- If the drug effect is mediated by multiple active moieties, the relevance of genetic variation is diminished. For example, if the product of the drug–enzyme reaction is an active metabolite with activity similar to that of the parent molecule, an altered ratio of parent drug to metabolite (through polymorphism) may have little clinical effect.
- If a drug is metabolized or eliminated by multiple pathways, total ablation of one pathway (as in a PM) may result in minimal alteration of overall drug concentrations and hence have minimal effect.
- Sometimes enantiomers have different pharmacological activity and pathways of elimination. This is called stereoselective metabolism.
- Phenocopying is the conversion of a patient from a phenotypic normal metabolizer into a slow metabolizer as a result of inhibition of the enzyme by another drug or by itself ("autophenocopying"). This can result in a reduction in a trait's variation in a population. For example, all patients chronically taking a drug are poor metabolizers either because of their genotype or because of autophenocopying. Phenocopying sometimes explains why differences between PMs and EMs are less at steady state than after single doses.

17.3.1 CYP450 Enzymes

The major causes of interindividual and intraindividual variability in CYP activity are environmental factors (inducers and inhibitors), biological factors (gender, disease, and circadian rhythms), and genetic polymorphisms in CYP450 genes and their regulators. There are large variations between individual CYP450 isoforms in terms of their susceptibility to these mechanisms [35,36]. Overall, 57 CYP450 genes and 58 pseudogenes have been identified, 42 of which play a role in the metabolism of both exogenous xenobiotics and endogenous substances (e.g., steroids and prostaglandins), and 15 of which are involved in the metabolism of drugs in humans [35,36]. CYP450 genes are highly polymorphic and can exhibit clinically significant genetic polymorphisms. In general, CYP3A4/5, CYP2D6, CYP2C9, CYP2C19, CYP2A6, CYP2B6, and CYP2C8 are the most important and most studied metabolic enzymes. The variations in CYP450 genes—deletions, missense mutations, deleterious mutations (creating splicing defects or premature stop codons), and duplications—can result in abolished, reduced, normal, or enhanced enzyme activity [21,35,37,38].

17.3.1.1 CYP2D6

CYP2D6 is a member of the cytochrome P450 superfamily involved in metabolizing and eliminating many prescribed medications; it accounts for approximately 2% of total hepatic CYP450 content. The *CYP2D6* gene is highly polymorphic, with at least 100 genetic variants and 120 alleles identified. Additionally, its variants are the best characterized among all CYP450 variants, and the distributions of these alleles exhibit notable interethnic differences. (See Chapter 16.) Approximately 7–10% of Caucasians and 1% of Chinese, Japanese, and Koreans are PMs of CYP2D6 [34,39,40].

In Caucasians, *CYP2D6**3, *4, and *5 produce inactive enzymes or no protein products and are the variants most commonly implicated in the PM phenotype. *CYP2D6**4 is the most common variant allele in Caucasians (allele frequency ~21%), but it is virtually absent in Chinese, although overall CYP2D6 activity is lower in Chinese than in Caucasians as a result of the high allele frequency of *CYP2D6**10 (~50%), which is largely absent in Caucasians. Gene duplication occurs in ~1–7% of Caucasians and 29% of black Ethiopians; also, it is predictive of an UM phenotype [34]. This variant produces an unstable enzyme with reduced (but not absent) ability to metabolize substrate drugs.

CYP2D6 is responsible for the metabolism of almost 20–30% of drugs and is the most widely studied enzyme in relation to polymorphisms. Examples include tricyclic antidepressants, selective serotonin reuptake inhibitors (SSRIs), antiarrhythmics, beta-blockers, neuroleptics, opioid analgesics, antiemetics, and anticancer drugs. CYP2D6 is not inducible, and therefore variations in the enzyme expression and activity are largely attributable to genetic polymorphisms [21,41]. Powerful CYP2D6 inhibitors have been shown to significantly decrease EM metabolic capacity; therefore, EM subjects may seem to be PMs during the coexisting administration of potent inhibitors (e.g., fluoxetine and paroxetine) [39].

Recently, ADRs have been reported in UMs, mainly of a 10- to 30-fold increase in drug metabolite concentrations. For example, codeine is converted through CYP2D6 to morphine, which is more pharmacologically active. UMs receiving standard doses of codeine have been reported to display symptoms of narcotic overdose significantly related to higher morphine concentrations [39,42]. Perhexiline is an antianginal drug that is almost entirely metabolized by CYP2D6 to hydroxyperhexiline (inactive). PMs have trough concentrations up to 6-fold higher than EMs after a single dose, with evidence of saturable metabolism. The major toxicity of perhexiline is hepatotoxicity and peripheral neuropathy, which are concentration-dependent. Therapeutic drug monitoring has assisted dosing, with a suggested range of 0.15–0.6 mg/l, supported by both concentration-dependent efficacy and toxicity [43]. This supports a case for prospective genotyping.

In most individuals, only a small fraction (~10%) of codeine is metabolized to morphine via CYP2D6, with most of that fraction glucuronidated to codeine-6-glucuronide and the remainder metabolized by CYP3A4 to norcodeine. The AUC of codeine is similar in PMs and EMs, whereas morphine is virtually undetectable in PMs as well as EMs taking quinidine (phenocopying). Clinical studies in volunteers generally support the lack of analgesia in PMs, which is consistent with the belief that morphine is the key metabolite responsible for codeine's antinociceptive effects. Theoretically, UMs may convert codeine to morphine more rapidly, resulting in increased opioid effects for a given dose. However, studies to date have been with volunteers and not the target population.

Overall there is a strong argument for a gene–concentration effect resulting from the failure of prodrug conversion to morphine in PMs and phenocopied EMs. As far as the gene–effect relationship is concerned, there seems to be a predictable failure of analgesia in healthy volunteers but a less clear relationship with adverse effects. Although nonresponse to codeine may be explained by PM status, it also may be due to other factors, including phenocopying by a CYP2D6 inhibitor such as paroxetine [34].

The beta-1-selective metoprolol appears to have both consistent gene–concentration and gene–effect relationships in healthy volunteers, suggesting that dose reduction to ~25% should occur in PMs or those phenocopied by other drugs [34]. Metoprolol is administered as a racemate, with S-metoprolol thought to produce most of the β-blockade. The main metabolite, O-desmethylmetoprolol (essentially inactive), accounts for approximately two-thirds of the metabolism and occurs via various pathways, including CYP2D6 (mainly R-metoprolol); another pathway, to α-hydroxymetoprolol, accounts for ~10% of the dose in EMs and seems to be under CYP2D6 control because very little is produced in PMs. AUCs of metoprolol are 4- to 6-fold higher in PMs than in EMs after one dose and 3- to 4-fold higher after repeated dosing. UMs achieve metoprolol concentrations that are half those observed in EMs. The strong CYP2D6 inhibitor, paroxetine, increases the mean AUCs of S- and R-metoprolol by 5- and 8-fold, respectively.

Enhanced or prolonged β-blockade has been observed in both PMs and EMs receiving CYP2D6 inhibitors [44,45]. Three recent prospective studies failed to show a relationship between CYP2D6 and adverse effects with metoprolol [46–48]. The high therapeutic index of the drug and the fact that the effects of excessive β-blockade are usually easy to detect clinically (e.g., bradycardia) lessen the need for genotype testing. The long history of metoprolol use without genotyping or phenotyping suggests that these tests are unlikely to happen in practice.

17.3.1.2 CYP2C9

CYP2C9 is the most abundant of the CYP2C enzymes and accounts for about 30% of total hepatic CYPs. It plays an important role in the metabolism of about 10% of the drugs available on the market, including nonsteroidal anti-inflammatories (e.g., ibuprofen), antiepileptics (e.g., phenytoin), oral anticoagulants (e.g., warfarin), and antihypertensives (e.g., losartan). The human *CYP2C9* and *CYP2C19* genes are highly homologous at the nucleotide level [39,49], and more than 30 variants of *CYP2C9* have been identified. Generally, *CYP2C9* polymorphisms encode proteins with a loss of catalytic function, with the extent this reduction often being substrate-dependent.

Many variants have been associated with reduced enzyme activity (see http://www.cypalleles.ki.se), with *CYP2C9**3 and, to a lesser extent, *2 having the most clinical relevance. *In vitro* studies show that **3* is associated with a lower intrinsic clearance of substrate drugs than is *2. The effects of *CYP2C9**2 are more substrate specific (e.g., warfarin and phenytoin), whereas *CYP2C9**3 displays reduced catalytic activity toward the majority of

CYP2C9 substrates. The clinical importance of *CYP2C9* polymorphisms is demonstrated by the dose adjustment of an oral anticoagulant, warfarin, based on the *CYP2C9* genotype[4,39,49,50].

At least one *CYP2C9**2 or *3 allele is carried by ~20% and 12% of Caucasians, respectively, with ~2.5% being homozygotes for *2 or *3, or for compound heterozygotes for both alleles. The remaining two-thirds of the Caucasian population are wild types and have normal enzyme activity. The small proportion of individuals (~0.4% of Caucasians) homozygous for *CYP2C9**3 have the lowest ability to metabolize substrate drugs. The *CYP2C9**2/*2 and *2/*3 genotypes may also cause important reductions in the metabolism of some drugs (e.g., phenytoin). *CYP2C9**2 and *3 are rare in African American and Asian populations, with the wild type making up more than 95% of these populations [51].

The three oral coumarin anticoagulants—warfarin, acenocoumarol, and phenprocoumon—exist as S- and R-enantiomers. The S-enantiomers are CYP2C9 substrates and are responsible for most of the effects of warfarin and phenoprocoumon. In contrast, although S- and R-acenocoumarol have comparable activities, rapid elimination of the S-enantiomer means that R-acenocoumarol produces most of the anticoagulant effect. All three oral anticoagulants have low therapeutic indices, and the dose required to produce a target prothrombin time is largely unpredictable [34] (see Chapter 24). CYP2C9 is important for the metabolism of this class, with *CYP2C9**3 (but not *CYP2C9**2) being clearly implicated in impaired clearance of tolbutamide, glyburide, and glimepiride [52–54].

Losartan is metabolized by CYP2C9 via an aldehyde intermediate (E-3179) to E-3174, the predominant active moiety. E-3174 is at least 10-fold more potent than losartan at the AT1 receptor. Furthermore, although only 14% of losartan is metabolized to E-3174, the AUC of the latter is 4-fold to 8-fold higher than that of the parent and is thought to be responsible for most of the activity [55]. This finding suggests that individuals with *CYP2C9* variants might exhibit a reduced losartan response. Only one of the three studies in healthy volunteers that included blood pressure assessments reported a significant influence of CYP2C9; they reported reduced response among those with the *1/*3 genotype [56–58]. Studies on the impact of genotype on AUC have been inconclusive, and it seems unlikely that *CYP2C9* variants significantly affect parent losartan concentrations since production of E-3174 constitutes a quantitatively minor route of elimination [56–59].

Phenytoin is a commonly used antiepileptic drug despite its complex nonlinear pharmacokinetics and low therapeutic index. It is primarily (80–90%) eliminated via 4′-hydroxylation to 5-(4p-hydroxyphenyl)-5-phenylhydantoin (HPPH), largely via CYP2C9, which preferentially produces the S-enantiomer of HPPH. The presence of at least one *CYP2C9**2 or *3 allele is associated with one-third lower mean dose requirements (199 versus 314 mg/day, respectively). Furthermore, a "gene-dose" effect seems to exist, with dose requirements of 314, 193, 202, 217, and 150 mg/day for the *CYP2C9**1/*1, *1/*2, *1/*3, *2/*2, and *2/*3 genotypes, respectively [60].

In a single-dose study in healthy volunteers 30% lower concentrations were seen in wild-type individuals compared with carriers of *CYP2C9**2 or *3 alleles. *CYP2C9**2 and *3 were suggested as accounting for 31% of the variation in phenytoin concentrations taken 12 h after a single 300 mg dose [61]. As central nervous system toxicity (e.g., ataxia and nystagmus) is closely related to concentration, it is likely that individuals with *CYP2C9* variants are predisposed to these effects. There is some evidence for an association between *CYP2C9**1 and *3 genotype and cutaneous reactions [62]. A reasonable case for genotyping in relation to phenytoin to guide the initial choice of a maintenance dose can be made on the basis of a strong gene-(dose) concentration effect, a moderate gene-effect relationship, a strong concentration-effect relationship for both desired and adverse effects, and a low therapeutic index. Points against genotyping include the fact that an alternative pathway (CYP2C19) exists for phenytoin metabolism. Therapeutic drug monitoring continues to be necessary, and clinicians have long experience with phenytoin.

17.3.1.3 CYP2C19

CYP2C19 accounts for about 3% of total hepatic CYPs. The ability of individuals to metabolize S-mephenytoin has enabled them to be classified as PMs or EMs. PM phenotypes for CYP2C19 are common (20%) among Asians and rare (3–5%) in European-derived populations. CYP2C19 catalyzes the metabolism of several drugs, including proton pump inhibitors (PPIs) (e.g., omeprazole, lansoprazole, pantoprazole), antidepressants (e.g., citalopram and amitriptyline), antiplatelet drugs (e.g., clopidogrel), antifungals (e.g., voriconazole), and anticancer compounds (e.g., cyclophosphamide). Seven variants (*2–*8) in the *CYP2C19* gene have now been associated with reduced enzyme activity *in vivo*, largely due to production of inactive enzyme protein.

A novel *CYP2C19* variant (*CYP2C19**17) that may produce an ultrarapid metabolizer phenotype was recently identified [63]. A splice site mutation, *CYP2C19**2 (rs4244285, 19154G >A), and a premature stop codon, *CYP2C19**3 (rs4986893, 17948G >A) represent the two most predominant null alleles [4,39,64]. Genotyping for *CYP2C19**2 and *3 identifies most PMs in African American and Chinese populations, while genotyping for *CYP2C19**2 identifies 70–85% of variant reduced-activity alleles and *CYP2C19**2–*8 identifies more than 99% of PMs [65].

CYP2C19 plays an important role in proton-pump inhibitor (PPI) therapy for peptic ulcers and gastroesophageal reflux disease. In EMs, it is responsible for >80% of the metabolism of omeprazole, lansoprazole, and pantoprazole [66], the remainder with CYP3A4 [39,64]. The metabolites produced are inactive. A fourth PPI, rabeprazole, may be less reliant on CYP2C19 as it undergoes nonenzymatic conversion to rabeprazole thioether. S-Omeprazole (esomeprazole) was recently marketed as an individual entity to exploit its reduced variation in *CYP2C19* genotype pharmacokinetics compared with the racemate or with R-omeprazole [67].

The AUCs of both omeprazole and lansoprazole are 4- to 15-fold higher in PMs than in homozygous EMs; with multiple dosing, the increase in AUC of omeprazole (but not of lansoprazole or pantoprazole) decreases to ~2-fold in EMs because of inhibition of its own metabolism by CYP2C19 [68]. This does not occur in PMs who lack a functioning CYP2C19 enzyme to inhibit. The AUCs of rabeprazole are also increased but less markedly in CYP2C19 deficiency. Individuals with CYP2C19 deficiency have superior acid suppression with conventional doses of omeprazole and lansoprazole [69], and increasing the lansoprazole dose from 30 mg once daily to four times daily in homozygous EMs leads to an increase in mean 24-h intragastric pH from 4.5 to 7.0 [70].

Overall, omeprazole- and lansoprazole-based regimens produce lower eradication rates in homozygous EMs than in heterozygous EMs or PMs [71]. If patients are confirmed as being PMs, dual therapy with PPI and amoxicillin may be appropriate, as the eradication rate is likely to be high (>90%) [72]. The PPIs are an exceptionally well-tolerated class of drugs, and there seems to be no clear evidence of increased toxicity in PMs despite a markedly elevated AUC. However, individuals with a CYP2C19 deficiency may be predisposed to vitamin B12 deficiency during long-term use of this class [73].

17.3.1.4 CYP2A6

CYP2A6 is a member of the cytochrome P450 superfamily and is expressed predominantly in hepatic cells, with some expression in specialized extrahepatic cell types. Compared with CYP2D6 and CYP3A4, relatively few clinically used drugs are metabolized to a significant point by CYP2A6. Its substrates include coumarin, halothane, methoxyflurane, valproic acid, disulfiram, and losigamone [74]. A number of significant variations in CYP2A6 have been identified, including *CYP2A6**2 (rs1801272, 479T >A), *CYP2A6**4 (gene deletion), *CYP2A6**5 (rs5031017, 1436G >T), and *CYP2A6**20 (rs28399444, frame shift). These polymorphisms are associated with abolished enzyme activity and have different allele frequencies among ethnic groups. The prevalence of CYP2A6 PMs in the Caucasian population is ≤1% [39].

A deletion of the *CYP2A6* gene is common in the Asian population and accounts for significant differences in PMs compared with the Caucasian population. Since CYP2A6 is the high-affinity metabolizer of both nicotine and its oxidized metabolite, cotinine, *CYP2A6* variants have mainly been studied for treating tobacco abuse. Studies have revealed that the kinetics of nicotine metabolism are different in individuals carrying the variant *CYP2A6* alleles. For example, in three studies, smokers who carried the *CYP2A6* variants smoked fewer cigarettes and were more likely to quit smoking. These results reflect the possibility of increased nicotine concentrations and, subsequently, increased nicotine tolerance and ADRs from nicotine, especially in CYP2A6 PMs [39,74,75].

17.3.1.5 CYP2B6

CYP2B6 accounts for 6–10% of the total CYP contents in the liver, with a substantial (>100-fold) variation in expression between individuals [76]; it is involved in the metabolism of an increasing number of clinically important drugs (~8% of those on the market) [39,77], including bupropion, cyclophosphamide, efavirenz, and methadone. It also metabolizes some procarcinogens (e.g., 6-aminochrysene) and drugs of abuse (e.g., N-methyl-3,4-methylenedioxyamphetamine, "ecstasy"). CYP2B6 is subject to inhibition and induction by drugs such as clopidogrel and phenobarbital respectively [78,79]. A number of *CYP2B6* variants have been identified. *CYP2B6**6 (rs3745274, 516G >T and rs2279343, 785A >G) is the most common polymorphism. It commonly occurs in Caucasians and Asians, while *CYP2B6**16 (rs2279343, 785A >G) and *CYP2B6**18 (rs28399499, 21011T >C) are common in Africans [39].

CYP2B6 is the main catalyst of efavirenz metabolism (to its inactive 8-OH metabolite); therefore, its polymorphisms may have major implications for the PKs and toxicity of this drug, which at present is recommended as an option in first-line combination therapy for HIV infections [77]. Individuals homozygous for the 516T variant or the *CYP2B6**6 allele (516G >T and 785A >G) may have two to threefold higher efavirenz concentrations and may be predisposed to side effects. Studies have shown that when the efavirenz dose is decreased, the plasma concentrations of the drug decreases proportionally. It is reported that patients who were *6/*6 homozygous and *6/*26 and *1/*26 heterozygous had lower plasma concentrations of efavirenz and less frequent CNS effects when the dose was decreased [39,80].

17.3.1.6 CYP2C8

CYP2C8 accounts for approximately 7% of total hepatic content and plays a vital role in the metabolism of pioglitazone, amiodarone, paclitaxel, chloroquine, verapamil, and ibuprofen. It also plays a secondary role in the metabolism of fluvastatin, amitriptyline, diclofenac, omeprazole, and carbamazepine. The metabolism of paclitaxel to 6α-hydroxypaclitaxel, which is essentially inactive, has been used as an index of CYP2C8 activity *in vitro* [81]. The most common *CYP2C8* variants are *CYP2C8**2 (rs11572103, 11054A >T), *CYP2C8**3 (rs11572080, 2130G >A and rs10509681, 30411A >G), and *CYP2C8**4 (I264M substitution); they lead to decreased enzyme activity [21,39,82]. CYP2C8 is primarily responsible for the hydroxylation of R-ibuprofen; CYP2C9, for S-ibuprofen. The mean AUC after a single dose of 400 mg of racemic ibuprofen was reported to be 2.2- and 8.7-fold higher in individuals with the *CYP2C8**1/*3 or *3/*3 versus *1/*1 genotype [83].

*CYP2C9**2 has been associated with altered ibuprofen pharmacokinetics only when it is coinherited with *CYP2C8**3 (linkage disequilibrium) [83]. Several studies suggest that reduced paclitaxel metabolism may occur with the *CYP2C8**2 and *3 alleles, with the latter producing the greatest reduction [84]. *In vitro* data suggest that CYP2C8 and CYP3A4 are responsible for inactivating repaglinide, a nonsulfonylurea insulin secretagogue. CYP3A4 inhibitors such as ketoconazole and clarithromycin increase repaglinide concentrations by small amounts (15% and 40%, respectively) compared with a 5.5- to 15.5-fold increase with the CYP2C8 inhibitor gemfibrozil [85].

*CYP2C8**2 is more common in Africans, while *CYP2C8**3 and *4 are common in Caucasians. Both CYP2C8 and CYP2C9 play an important role in ibuprofen metabolism, and it has been found that CYP2C8 plays a key role in the hydroxylation of the (R)-enantiomer. For this reason, ibuprofen disposition has been used as a guide for CYP2C8 activity. When the effects of the different genotypes on ibuprofen PKs were investigated, *CYP2C8* polymorphisms were found to be common [82,86].

Studies have shown that the presence of the *CYP2C8**3 allele causes a significant effect on the disposition of ibuprofen, which suggests that a considerable proportion of Caucasian subjects may show alterations in the disposition of drugs that are CYP2C8 substrates. Moreover, in one study *CYP2C8**3 carriers who received R-ibuprofen had lower ibuprofen CLs; the *CYP2C8**3 heterozygotes who were treated with the antidiabetic repaglinide had a higher metabolism of the drug when compared to *CYP2C8**1 and *4 carriers [39,82,86].

17.3.1.7 CYP3A4/5

The CYP3A locus comprises four genes that code for the functional enzymes: CYP3A4, CYP3A5, CYP3A7, and CYP3A43. Of these CYP3A7 is primarily fetal and CYP3A43 is minimally expressed and has low functional activity. CYP3A4/5 is responsible for metabolizing the largest number of prescribed drugs (more than 50%), including cyclosporin, indinavir, nifedipine and simvastatin. These genes are often considered collectively as "CYP3A" because of their promiscuous substrate specificity and the difficulty discerning the relative role of each isoform in drug metabolism [34]. The overall activity of CYP3A is unimodally distributed, exhibits wide interindividual variability (>10-fold), and is highly susceptible to the effects of enzyme inhibitors and inducers [87].

More than 20 variants in the coding region of CYP3A4 have been described [39], and a number of them have altered enzyme activities, ranging from a modest to a significant loss in catalytic efficiency; however, their clinical significance is uncertain. *CYP3A4**2 (rs55785340, 664T >C) and *CYP3A4**7 (rs56324128,167G >A) have higher frequencies in Caucasian populations, whereas *CYP3A4**16 (rs12721627, 554C >G) and *CYP3A4**18 (rs28371759, 878T >C) have higher frequencies in Asian populations [4].

A single point mutation (6986A >G) in intron 3 of CYP3A5 (designated CYP3A5*3) produces a truncated and nonfunctional protein. In contrast to the unimodal distribution of CYP3A taken as a whole, CYP3A5 exhibits a bimodal distribution that can be predicted by the presence or absence of this allele. Individuals homozygous for CYP3A5*3 produce little CYP3A5 enzyme ("low expressors"), whereas the remainder have at least one wild-type (CYP3A5*1) allele and express a large amount of CYP3A5 ("high expressors").

The PKs of the immunosuppressive agent tacrolimus are mainly dependent on the *CYP3A5* genotype, and in terms of the dose required to reach target trough blood concentrations (C0). Tacrolimus C0 and its dose requirements are related to *CYP3A5* polymorphisms; for example, individuals who carry at least one *CYP3A5**1 allele have a functional CYP3A5 and consequently need a higher dose of tacrolimus to reach the targeted blood concentration. CYP3A5 expressors experience a delay in reaching target concentrations [4,88]. It is estimated that 10–20% of Caucasians, 40–50% of East Asians, 60–70% of Hispanics, and >80% of African Americans are high expressors of CYP3A5 [89].

17.3.2 Non-CYP450 Enzymes

Here we discuss the variations in the non-CYP450 enzymes: UGT, TPMT, NAT, and GST.

17.3.2.1 UDP Glucuronosyltransferases

UDP Glucuronosyltransferases (UGTs) are expressed in a tissue-specific and frequently inducible way in most human tissues; their highest concentration is found in the GI tract and liver. UGTs contribute to about 35% of phase II drug metabolism and are involved in the glucuronidation of many endogenous compounds and xenobiotics [39]. Nineteen human genes are encoded by UGT proteins: 9 by *UGT1* and 10 by *UGT2*. While both types are involved in the process of drug metabolism, UGT2 seems to have greater specificity for the glucuronidation of endogenous substances (e.g., steroids). *UGT1A1* expression plays a vital role in drug metabolism because the glucuronidation of bilirubin by UGT1A1 is the rate-limiting step that assures effective bilirubin conjugation levels. This rate can be influenced by both genetic variation and drug treatments. If a drug undergoes selective metabolism by UGT1A1, competition for its metabolism with bilirubin, glucuronidation occurs, resulting in marked hyperbilirubinemia in addition to decreased clearance of the metabolized drug [4,20,90].

UGT1A1 recently gained recognition as the first pharmacogenetic test to achieve FDA approval for use in conjunction with a specific drug (irinotecan). Irinotecan is a prodrug that is metabolized by carboxylesterases to the active topoisomerase inhibitor 7-ethyl-10-hydroxycamptothecin (SN-38) and by CYP3A4 to inactive metabolites. Thereafter, SN-38 is glucuronidated by UGT1A1, with the resultant SN-38 glucuronide excreted into the intestine via bile [91].

The activity of UGT1A1 varies widely, with an *in vitro* study demonstrating a 17-fold variation in SN-38 glucuronidation. *UGT1A1**28 is the variant most frequently implicated in defective SN-38 glucuronidation and involves an extra thymine-adenine (TA) repeat in the TATA section of the *UGT1A1* promoter (i.e., $(TA)_7TAA$ instead of $(TA)_6TAA$ in the wild type). It is also the primary cause of Gilbert's syndrome [92]. This variant occurs commonly, with the homozygous genotype found in 5–15% of Europeans, 10–25% of Africans and South Asians, and 1–5% of Southeast Asians and Pacific Islanders. There is an increased risk of severe neutropenia and diarrhea with irinotecan in homozygotes for *UGT1A1**28 compared with homozygotes for *UGT1A1**1 [93].

Gilbert's syndrome can be caused by a number of genetic changes, and patients with it may experience adverse drug reactions (ADRs) because of reduced capacity to metabolize drugs by UGT1A1. *UGT1A1**37 has an 8-TA insertion in the promoter and results in more decreased promoter activity than that of the *UGT1A1**28 allele. *UGT1A1**36 (rs8175347) has only a 5-TA insertion in the promoter and is associated with increased enzyme activity and a decreased risk of neonatal hyperbilirubinemia [20,90].

17.3.2.2 Thiopurine S-methyltransferase

Thiopurine S-methyltransferase (TPMT) catalyzes the S-methylation of 6-mercaptopurine (6-MP), azathioprine (AZA), and thioguanine to inactivate these drugs, which are used for treating leukemia and autoimmune diseases. It is a cytosolic enzyme found in many tissues, with activity most commonly determined in red blood cells. In Caucasians, a trimodal distribution exists, with 0.3–0.6% having low or undetectable activity, 10% having intermediate activity, and the remaining 90% having high (normal) activity [94]. The *TPMT* gene exhibits significant genetic polymorphisms across all of the ethnic groups studied, and TMPT activity is highly variable, as well as polymorphic, in all large populations evaluated up to the present time. About 90% of Caucasians inherit high enzyme activity, 10% inherit intermediate activity (heterozygous), and 0.3% inherit low or no activity. A number of clinically significant SNPs have been identified for *TMPT* that can alter its activity, and since methylation is involved in both the activation and the metabolism of mercaptopurine, altered enzyme activity affects the concentration of both active and toxic metabolites [4,21,39].

The genetic polymorphisms of *TMPT* have the strongest case for prospective pharmacogenetic testing. Generally, thiopurine agents have a narrow therapeutic index, with life-threatening myelosuppression being a major concern. Patients with polymorphic TMPT frequently require a significant dose reduction in order to prevent toxicity. More than 20 variants of the *TPMT* gene have been identified; however, the most studied are *TPMT**2, *3A, *3B, and *3C, which are responsible for most cases of TPMT deficiency. These three variants account for 80–95% of intermediate and poor metabolizers [4,21,39]. *3A is the most common variant in Caucasians; *3C is the most common in Africans and Southeast Asians. Homozygotes for a variant allele (e.g., *TPMT**3A/*3A) have negligible TPMT activity, whereas heterozygotes (e.g., *TPMT**1/*3A—intermediate activity) have an average activity that is approximately half that observed in homozygotes for wild-type alleles (*TPMT**1/*1, normal/high activity). Genotyping is reasonably accurate in predicting TPMT phenotype, which is defined as low, intermediate, or normal/high TPMT activity [95].

The therapeutic niche of TPMT testing relates to its ability to identify prospectively the small proportion of patients (0.3–0.6%) with enzyme deficiency who will almost certainly develop life-threatening myelosuppression if standard doses are used. Although it is clear that TPMT deficiency (particularly complete deficiency) predisposes to myelotoxicity, approximately 75% of cases of bone marrow depression occur in those without mutations (i.e., presumed normal enzyme activity) [96]. This indicates that TPMT metabolizer status is a useful adjunct to (but not replacement for) regular blood-count monitoring. The clinical usefulness of prospective determination of TPMT metabolizer status is supported by pharmacoeconomic studies [97].

17.3.2.3 Dihydropyrimidine Dehydrogenase

Dihydropyrimidine dehydrogenase (DPD), an enzyme encoded by the DPYD gene, metabolizes two endogenous pyrimidines—thymine and uracil—and facilitates the metabolism of the pyrimidine analog 5-fluorouracil (5-FU). DPD activity in peripheral blood mononuclear (PBM) cells has been used as a surrogate for total body DPD activity and varies up to 20-fold. DPD activity is usually described as normally distributed, perhaps reflecting the absence of cases of absolute deficiency in the studies [98]. DPD polymorphisms result in DPD-deficient phenotypes with a total frequency of about 3–5%. DPD variants, such as *DPYD**13 (rs55886062, 1679T >G), *DPYD**9A (rs1801265, 85T >C), *DPYD**2A (rs3918290, IVS14+1G >A), and 2846A >T (rs67376798, D949V), are among the identified SNPs associated with grade 3 and grade 4 toxicities in 5-FU–treated patients. For example, *DPYD**2A homozygous patients have a complete absence of normal DPD activity, whereas heterozygous carriers have a 50% absence of the enzyme activity, resulting in 5-FU accumulation and significant life-threatening 5-FU–related toxicities [39].

Moreover, *DPYD**9A is a common missense mutation with a C29R substitution; however, its association with reduced DPD activity is still debatable. For example, two different studies reported that cancer patients who were *DPYD**9A heterozygous had severe 5-FU toxicity in. Together, DPYD*2A and *DPYD**9A polymorphisms have a high concordance with 5-FU toxicity but a low concordance with enzyme activity. Determination of PBM cell DPD activity (i.e., phenotype) may identify up to 60% of patients who may develop severe toxicity, whereas screening solely for the *DPYD**2A allele (genotype) identifies approximately 25% of these patients. Identification of a patient with absolute deficiency would allow selection of alternative chemotherapy, whereas those with partial deficiency can be treated with a lower dose of 5-FU. The increasing availability of DPD inhibitors (eniluracil) may make assessment of metabolizer status redundant.

17.3.2.4 N-acetyltransferases

Human N-acetyltransferases (NATs) catalyze the acetylation of aromatic amines and hydrazines. There are two functional NAT genes in humans, *NAT-1* and *NAT-2*, which carry functional polymorphisms that influence enzyme activity [4]. More than 25 allelic variants of these functional genes have been identified. Based on the level of NAT activity, patients can be classified into two phenotypes: fast acetylator (FA—wild-type NAT acetylation activity) and slow acetylator (SA—reduced NAT enzyme activity). NAT-1 activity is usually constant, whereas NAT-2 polymorphisms result in individual differences in the rate at which drugs are acetylated. While several mutations have been identified in the *NAT1* and *NAT2* genes, the frequency of the slow acetylation patterns is attributed mostly to polymorphisms in the *NAT2* gene [4,20,21].

Altogether, *NAT2* polymorphisms result in more than 10 *NAT2* alleles, and the variant alleles that account for the majority of SA phenotypes include *NAT2**5 (rs1801280, 341T >C), which results in an I114T substitution; *NAT2**6 (rs1799930, 590G >A), which causes a R197Q substitution; and *NAT2**7 (rs1799931, 857G >A), which corresponds to a G186E substitution [39]. A number of compounds have been used as a probe for NAT2, and a good genotype–phenotype correlation is generally observed (>90%) with caffeine as a probe [99].

Slow acetylators are found in approximately 50% of European- and African-derived populations, but they are less common in Asians. Individuals phenotyped as slow acetylators are likely to have two slow activity alleles, whereas those phenotyped as fast acetylators can have one or two high-activity alleles (most likely *NAT2**4, which is considered to be the wild type). In Caucasians, most of the fast acetylators are heterozygotes for slow and fast alleles, whereas individuals homozygous for fast alleles (e.g., *NAT2**4/*4) comprise ~5% of the population [99].

In contrast, ~30% of Chinese individuals are homozygous for rapid alleles, 45% are heterozygous, and 25% are homozygous for slow alleles [100]. *NAT2* polymorphisms and their association with the slow acetylation of isoniazid (an antituberculosis drug) were one of the first fully characterized genotypes shown to influence drug metabolism, which links the pharmacogenetic (PG) phenotype to a genetic polymorphism [20] (see Chapter 36). Since NAT1 and NAT2 catalyze the bioactivation (via O-acetylation) of aromatic and heterocyclic amine carcinogens, genetic variations in the *NAT1/2* genes may alter the cancer risk associated with exposure to these carcinogens. *NAT2* polymorphisms are related to individual susceptibility to particular cancers caused by industrial chemicals (e.g., α- and β-naphthylamine); for example, individuals with poor acetylator phenotypes have increased risks of lung, bladder, and gastric cancers if exposed to carcinogenic arylamines for a long period of time [4,21].

17.3.2.5 Glutathione S-transferases

The superfamily of human glutathione S-transferases (GSTs) facilitates the conjugation of glutathione (GSH) to different endogenous metabolites and xenobiotics. GSTs include 10 members, which are divided into 3 main groups: cytosolic/nuclear, mitochondrial, and microsomal. Cytosolic GSTs are also divided into 7 classes: alpha, mu, omega, pi, sigma, theta, and zeta, the most important of these being *GSTM1*, *GSTT1*, *GSTP1*, and *GSTA1*. In addition to their enzymatic function, GSTs also act as regulators of many physiological processes, including cell signaling and growth factors and DNA replication [21].

The PKs of many drugs can be affected by *GST* polymorphisms, and since they decrease the activity of metabolizing enzymes, biologically active parent drugs or metabolites can accumulate and reach toxic levels. A number of known *GST* variants are well characterized in their influence on drug disposition. Two polymorphisms of the *GSTP1* gene have been identified: 1404A >G SNP (rs947894), with an I105V substitution, and 2294 C >T SNP (rs1799811), with an A114V substitution [39].

17.3.3 Drug Transporters

The movement of drugs across the cell membrane is through a combination of both passive diffusion and active transport, facilitated by certain drug uptake and efflux molecules. Drug transporters are membrane proteins that exist at various endothelial and epithelial barriers, including the BBB, intestinal epithelial cells, hepatocytes, and renal tubular cells. These transporters can significantly affect drug disposition. For example, the influx of a drug from the blood to the liver, where it is subsequently metabolized and excreted, may increase the rate of elimination. These proteins and the genes that encode them are essential to drug uptake, bioavailability, targeting, efficacy, toxicity, and clearance. Genetic variation in the genes encoding these transporters can result in variable expression levels and transport efficiency, which can have an impact on drug pharmacokinetics and response to treatment.

More than 300 human genes code transporters or transporter-related proteins, most of which work on endogenous substrates, although some also transport xenobiotics. Drug transporters can generally be classified into two groups: the efflux adenosine triphosphate-binding cassette (ABC), known as the multidrug resistance (MDR) family of transporters, and the uptake solute carrier single-level cell (SLC) family of transporters. SLCs mediate passive movement of solutes down their electrochemical gradient, while ABCs are active pumps fueled by ATP [39,101].

17.3.3.1 ABC Transporters

There are a total of 49 members of the human ABC transporter family, and they are grouped into seven subclasses or families (ABCA through ABCG) [102]. These transporters generally counteract uptake through the intestinal wall, efflux substrates out of tissues into systemic circulation, and eventually mediate the clearance of drugs from the body. Of all ABC transporters, the best known are ABCB1 (P-glycoprotein, Pgp, or MDR1), ABCC1/2 (multidrug resistance–related proteins 1/2, MRP1/2), ABBC2 (multidrug resistance, MRP2) and ABCG2 (breast cancer resistance protein, BCR). ABCB1 and ABCG2 are the most characterized [103]. They are expressed in the enterocytes, colon, intestinal epithelium, canalicular plasma membrane of hepatocytes, proximal renal tubule, hematopoietic stem cells, blood–brain barrier, heart, nerves, testes, and placenta. In all of these tissues, except the blood–brain barrier, they mediate efflux substrates out and into systemic circulation. ABCB1 in the choroid plexus, transports molecules from circulation into cerebrospinal fluid.

It is believed that the evolutionary role of ABC transporters is to limit the penetration of toxic molecules into critical organs, thereby protecting blood–tissue barriers [104,105]. ABCC1 is expressed ubiquitously and is localized to the basolateral, rather than apical, membranes of epithelial cells. Because of its basolateral localization, ABCC1 pumps drugs into the body rather than into the bile, urine, or intestine. For this reason, it is thought to serve mainly as a protective barrier in tissue epithelial cells rather than as a classic drug efflux pump. ABCC2 is functionally similar to ABCB1 and is expressed on the apical domain of epithelial cells in liver, intestine, and kidney [104,105].

ABCB1

ABCB1/MDR1 was the first recognized and most extensively studied ABC transporter encoding a polypeptide (P-glycoprotein, or P-gp). P-gp has two halves, each containing six hydrophobic transmembrane domains and an ATP-binding domain. The expression of *ABCB1* on the apical surface of enterocytes and the canalicular part of the hepatocyte has been shown to influence intestinal drug absorption and limit oral bioavailability of a wide variety of structurally diverse drugs. Additionally, it facilitates hepatobiliary excretion and renal excretion, and protects the brain and fetus from xenobiotics. ABCB1 overexpression in cancer cells is involved in multidrug resistance to chemotherapeutic agents [21,106,107].

ABCB1s are responsible for amphipathic lipid-soluble drugs. They transport a broad spectrum of structurally and functionally different drugs, such as anticancer agents (e.g., anthracyclines), antibiotics (e.g., erythromycin), immunosuppressants (e.g., cyclosporine), cardiac drugs (e.g., digoxin), calcium channel antagonists (CCB) (e.g., diltiazem), and HIV protease inhibitors (e.g., ritonavir).

The most common SNPs are the synonymous 1236C >T (rs1128503) and 3435C >T (rs1045642), and the nonsynonymous 2677G >T (rs2032583) [21,106], all three of which have allele frequencies that vary in different ethnic populations, with a higher frequency among Caucasians and Asians. The 3435C >T SNP has strong linkage disequilibrium with other SNPs in the *ABCB1* gene, creating common haplotypes consisting of 3435C >T combined with 2677G >T and/or 1236C >T. In general, SNP association studies on bioavailability and efficacy are inconclusive [105]. 3435C >T SNP was first shown to be associated with reduced serum digoxin concentrations, whereas it was also associated with higher plasma digoxin levels. The association was stronger when the *ABCB1* 2677G >T/A and 3435C >T polymorphisms were evaluated together as a haplotype [108].

Many investigators have now found similar associations between these polymorphisms and plasma concentrations of several other drugs, although these observations have not been consistently confirmed. More recent evidence suggests that polymorphic ABCB1 expression influences not only plasma pharmacokinetics but also the degree to which drugs are able to penetrate into tissues that express *ABCB1*. In a study of chronic myeloid leukemia patients treated with imatinib, patients with the 1236C >T had higher imatinib plasma concentrations and also showed an improved therapy response, whereas the presence of the wild-type 2677G variant worsened clinical response [109]. In a stable recombinant cell model, the anticancer drugs doxorubicin HCl, daunorubicin HCl, vinblastine sulfate, vincristine sulfate, taxanes (paclitaxel), and epipodophyllotoxin (etoposide, VP-16) exhibited selectively reduced Pgp-mediated resistance in 561A carriers [110].

In general, there is an overlap in substrate specificity and tissue distribution for ABCB1 and CYP3A4/5. Some drugs are dual substrates for both ABCB1 and OATP transporters (e.g., digoxin, fexofenadine, atenolol). Cyclosporine is not only transported by ABCB1, but is also metabolized by CYP3A4; consequently, the possible ABCB1 effects can be influenced by the activity of OATP transporters or CYP3A4 [21,39]. These may explain the conflicting results obtained with ABCB1 studies.

ABCC1 and ABCC2

ABCC1 and ABCC2 have overlapping substrate specificities, such as many anticancer agents (e.g., vincristine, doxorubicin), HIV protease inhibitors (e.g., ritonavir, saquinavir), and antibiotics (e.g., difloxacin, grepafloxacin). Both ABCC1 and ABCC2 need co-transport of reduced glutathione (GSH) to transport some of their substrates [12,21,111,112].

ABCC1 is ubiquitously expressed, whereas ABCC2 is mostly expressed in hepatocytes, renal proximal tubule cells, the intestine, and the brain. Genetic variants in *ABCC1* are relatively rare, and around 16 SNPs in both of its noncoding and coding regions have been determined to cause amino acid changes [21,112]. However, none of them significantly influence the expression level, indicating that the amino acid exchanges do not substantially affect the synthesis or stability of the protein. Gly671Val (rs45511401) in *ABCC1* and a haplotype of *ABCC2* were found to cause anthracycline-induced cardiotoxicity, as a consequence of the higher concentration of the drug, among non-Hodgkin lymphoma patients treated with doxorubicin [21].

In general, the influence of ABCC2-medicated transport on the pharmacokinetics of drug substrates in humans is not fully understood. Tumor cells often show an inducible expression of *ABCC2*, which contributes to drug resistance. Because of the transport of bile acids and glutathione from the hepatocytes into the bile duct, ABCC2 plays a physiologically important role in forming bile flow and potentially in detoxification by delivering glutathione for conjugation of xenobiotics.

The polymorphism −24C >T in the 5′-UTR is in strong linkage disequilibrium with 3972C >T. Renal allograft transplant recipients harboring the −24T allele show a decreased oral clearance for the immunosuppressant mycophenolic acid (MPA), which is the active metabolite of mycophenolate mofetil. In consequence, these patients are more protected from a decrease in MPA exposure but with a higher association to mild liver dysfunction [113]. High-dose methotrexate treatment in pediatric ALL induced a twofold higher AUC and a ninefold higher risk of intensification of folinate rescue in female patients carrying the −24 variant allele [114]. In non-small-cell lung cancer patients, *ABCC2*−24TT and 3972TT genotypes were associated with higher response rates and longer progression-free survival, which was sustained in haplotype analysis.

This suggests a more effective exposure to irinotecan [115]. A missense SNP 1249G >A (Val417Ile) is located in the substrate-binding region of the first transmembrane domain and is associated with lower oral bioavailability and increased residual clearance after intravenous administration of the beta-blocker talinolol, which indicates higher intestinal transporter activity [116]. It has also been found to be correlated with renal proximal tubulopathy after treatment with the HIV protease inhibitor tenovir disoproxil fumarate, probably because of a toxic concentration of the drug in renal tubular cells after being actively secreted into them from the blood by ABCC2 [117]. The silent polymorphism 1446C >G is associated with higher ABCC2 mRNA expression in the liver and with a decreased AUC and peak concentration (C_{max}) of the cholesterol-lowering drug pravastatin [118].

ABCG2

ABCG2—also known as breast cancer resistance protein (BCRP) or mitoxantrone-resistant protein (MXR) or placenta-specific ATP binding cassette transporter (ABCP)—is an ABC half-transporter with six transmembrane domains and only one ATP-binding domain. Overall, it is expressed at various sites, including the liver, small intestine, colon, placenta, lung, kidney, adrenal and sweat glands, and central nervous system vasculature.

More than 80 polymorphisms in the *ABCG2* gene have been identified in different ethnic populations, with a higher frequency among Asians and Caucasians. A number of *ABCG2* SNPs have been found to influence the function and/or expression of the encoded protein. For example, *ABCG2* variant alleles alter drug exposure by reducing the biliary excretion of diflomotecan and rosuvastatin, causing variations in drug effects [4,119].

The nonsynonymous variant, 421C >A SNP (rs2231142), affects the PK of many drugs, including irinotecan, rosuvastatin, sulfasalazine, and topotecan. It has been reported that the 421A variant exhibits 1.3-fold decrease in ATPase activity compared to the wild type, with elevated bioavailability of diflomotecan and topotecan. The AUC of rosuvastatin following a single oral administration has been shown to be greater in 421C >A heterozygotes and homozygotes, but the heterozygotes did not show any difference in the PK profile of irinotecan and its active metabolite, SN-38 [119]. The AUC and peak concentration of rosuvastatin increased 2.4-fold in healthy individuals with the homozygous 421AA genotype compared to that in individuals with the 421CC genotype, but did not affect the elimination of $t_{1/2}$, which signifies that 421C >A affects the intestinal absorption of rosuvastatin [120].

17.3.3.2 SLC Transporters

The SLC family comprises 360 members, subdivided into 47 subfamilies, that encode membrane proteins identified as passive transporters, ion-coupled transporters, and exchangers. The better-known SLC transporters are the organic anion–transporting polypeptides (OATPs) and the organic cation transporters (OCT).

Organic Anion–Transporting Polypeptides

Organic anion–transporting polypeptides (OATPs) have 12 transmembrane domains, with a large, highly conserved extracellular loop between the 9th and 10th transmembrane domains. *N*-glycosylation sites in extracellular loops 2 and 5 are consistent among the various members of the OATP family [121]. OATPs mediate the sodium-independent transport of a wide range of amphipathic organic compounds, including steroid conjugates, anionic oligopepties, thyroid hormones, bile salts, xenobiotics, and pharmaceuticals. A total of 11 OATPs have been identified and classified into 6 families, and in humans OATP1A2, OATP1B1, OATP1B3 and OATP2B1 are the best characterized for their roles in drug PK [4,39,122].

OATP1A2s encoded by the *SLCO1A2* gene are mainly expressed on the luminal membrane of small intestinal enterocytes and at the BBB, and play a role in the intestinal absorption and brain penetration of their substrates. OATP1A2 substrates include rosuvastatin, methotrexate. and D-penicillamine. A number of nonsynonymous polymorphisms have been identified in the *SLCO1A2* gene, some of which exhibit diminished *in vitro* transport activity toward OATP1A2 substrates [21,122,123].

The *SLCO1B1* gene encodes OATP1B1, also known as OATP-C, and is mainly expressed on the basolateral membrane of hepatocytes in the human liver. It plays a role in the hepatic uptake of substrate drugs, including statins, D-penicillamine, and rifampin [21,39]. The primary role of OATP1B1 is believed to be removal of substrates from the blood into the liver for subsequent elimination. More than 40 nonsynonymous variants have been identified in the *SLCO1B1* gene, some of them being associated with decreased transport activity and whose allele frequencies vary markedly across different ethnic populations. Functionally impaired OATP1B1 SNPs may decrease the uptake of substrate drugs into the hepatocytes, resulting in decreased biliary excretion or hepatic metabolism, which is a consequence of increased systemic exposure [21,120,124]. OATP1B1 is one of the mechanisms underlying both drug–drug interactions, due to competition at the transporter, and pharmacokinetic variation, due to genetic polymorphisms in the gene encoding the OATP1B1 protein *SLCO1B1*.

The most commonly studied OATP1B1 variants are the 521T >C SNP (rs4149056) and the 388A >G SNP (rs2306283), which are in linkage disequilibrium and exist together in several haplotypes. The variant allele 521T >C is more common among Caucasians (8–20%) and Asians (9–16%) than Africans [39,107]. It has consistently been found to cause a functional decrease in OATP1B1 activity based on altered *in vitro* transport of a number of substrates, including estrone-3-sulphate, estradiol-17β-D-glucuronide, pravastatin, atorvastatin, cerivastatin, rifampicin, and SN-38. Generally, there is a variant-dependent change in pharmacokinetics such that individuals homozygous for the 521T >C variant (CC) have the highest plasma concentrations, which is in line with *in vitro* data suggesting that this variant leads to a decrease in the function of OATP1B1 [111,125].

A genome-wide association study (GWAS) of 85 patients with myopathy and 90 matched controls from Study of the Effectiveness of Additional Reductions in Cholesterol and Homocysteine (SEARCH) identified an association with rs4149056 in the SLCO1B1 gene [126]. Following the accumulation of evidence on the SLCO1B1 521T >C polymorphism, the FDA announced an update to its simvastatin product-label recommendations in 2011.

SLCO2B1 encodes OATP2B1, which is also known as OATP-B, and shows substrate selectivity similar to that of OATP1B1. OATP2B1 is expressed in the luminal membrane of small-intestinal enterocytes, and have a role in drug absorption/disposition [39,122]. A number of OATP2B1 SNPs have been identified, including 1457C >T (rs2306168), 601G >A (rs35199625), 935G >A (rs12422149), and 43C >T (rs56837383). Patients who received montelukast, a leukotriene receptor antagonist, and who carry the 935A variant allele of the 935G >A SNP were reported to show lower plasma concentration of the drug as well as lower pharmacological response [39,122,123].

OATP1B3 (encoded by SLCO1B3) and OATP2B1 are mainly expressed on the sinusoidal membrane of hepatocytes, and mediate the hepatic uptake of their substrate drugs. OATP1B3 substrates include, digoxin, rifampin, methotrexate, D penicillamine, and cyclosporine. According to currently published findings, OATP1B3 appears to be unique in transporting digoxin and possibly also the taxanes docetaxel and paclitaxel [127]. Polymorphisms in SLCO1B3 include 334T >G (rs4149117) and 699G >A (rs7311358), both of which occur at a high frequency in Caucasian populations. While OAT1B3 mediates the hepatic uptake of many drugs, the role of its polymorphisms in drug disposition and response is not fully understood [111].

Organic Cation Transporters

Organic cation transporters (OCTs) belong to the solute carrier SLC22A family, which facilitates the intracellular uptake of a broad range of structurally diverse, small organic cations. Three OCT isoforms have been identified in humans: OCT1, which is encoded by *SLC22A1*; OCT2, which is encoded by *SLC22A2*; and OCT3, which is encoded by *SLC22A3*. OCT1 is primarily expressed in the basolateral membrane of hepatocytes, whereas OCT2 is expressed in the basolateral membrane of kidney proximal tubules. OCT1 and OCT2 substrates include metformin, cisplatin, imatinib, procainamide, citalopram, cimetidine, and quinidine. OCT3 is expressed in various tissues, including the placenta, heart, liver, and skeletal muscle. Additionally, OCT expression has been detected in several cancer cell lines and tumor tissue samples [21,39,107].

The 1393G >A polymorphism in *SLC22A1* was found to reduce the localization of OCT1 to the surface of the basolateral membrane of hepatocytes. Other variants, including 41C >T, 566C >T, 1201G >A, and 1256delATG (a deletion variant) have all been associated with decreased metformin uptake activity; increased AUC and Cmax; and the significantly decreased ability of metformin to lower glucose levels [111,128].

While several SNPs have been identified for the *SLC22A2* gene, the most important is the 808G >T SNP (rs316019). Since metformin's active renal secretion is through OCT2, any functional defect in OCT2 transport results in decreasing the drug's renal clearance. Homozygotes of low activity (808G >T SNP) have lower renal clearance and higher plasma concentrations of metformin compared to homozygous carriers of wild-type 270A [39,129]. Additionally, homozygous and heterozygous carriers of various haplotypes of low-activity alleles of some OCT1 variants—R61C (rs12208357), G401S (rs34130495), M420del (rs35191146), and G465R (rs34059508) among them—have shown significantly higher metformin renal clearance (20–30%).

The clinical use of cisplatin is related to dose-limiting nephrotoxicity, which occurs in one-third of patients regardless of intensive prophylactic measures. OCT2 has been involved in the cellular uptake of cisplatin, but its role in cisplatin-induced nephrotoxicity is not fully understood. Moreover, the nonsynonymous SNP 808G >T has been associated with reduced cisplatin-induced nephrotoxicity in patients. These results indicate the critical importance of OCT2 in PK and related cisplatin renal toxicity [39,129,130].

Previous studies showed that homozygous carriers of the low-activity OCT2 variant 270S have a significantly lower renal clearance and a higher plasma concentration of metformin than do homozygous carriers of the active variant 270A [21,129]. Additionally, homozygous and heterozygous carriers of various haplotypes of low-activity alleles of some OCT1 variants—including R61C (rs12208357), G401S (rs34130495), M420del (rs35191146), and G465R (rs34059508)—have shown significantly higher renal clearance of metformin PK (20–30%). In addition, low-function OCT1 variants (R61C, G401S, M420del, and G465R) have been associated with a significantly decreased glucose-lowering metformin response in healthy volunteers, probably by reducing metformin metabolism in hepatocytes—the major target sites of metformin's action. For OCT3, several synonymous SNPs have been identified in *SLC22A3*; however, their functional consequence has not been fully clarified and needs further study [21,39].

Organic Anion Transporters

Organic anion transporters (OATs) belong to the SLC22A family of solute-carriers, which facilitate the cellular uptake of a wide range of structurally diverse small hydrophilic organic anions. Several clinically important anionic drugs are OAT substrates, such as β-lactam antibiotics, diuretics, and nonsteroidal anti-inflammatories (NSAIDs). There are at least six OAT members (OAT1–6) [21], and four SLC22A genes that have been identified for encoding OATS: SLC22A6, which encodes OAT1; SLC22A7, which encodes OAT2; SLC22A8, which encodes OAT3; and SLC22A11, which encodes OAT4. In terms of tissue location, OAT1, OAT2, and OAT3 are expressed (mostly) in the basolateral membrane of the renal proximal tubules, mediating the uptake of drug substrates from the blood into the proximal tubule cells; OAT4, in contrast, is located at the apical side of the renal proximal tubule, functioning in the secretion of drug substrates into urine [21,107]. OAT1, OAT2, and OAT3 are responsible for the uptake of drugs into the tubular cells, and OAT4 mediates their secretion into the renal tubule [19,39]. Additionally, OAT1 is an example of transporter-related toxicity. Different studies have shown that OAT1substrates, such as cephaloridine and ß-lactam antibiotics, sometimes cause nephrotoxicity. OCT1 transports a number of drugs, such as metformin, which act partly through intracellular effects in hepatocytes [19].

Genetic variations in genes encoding OATs can contribute to interindividual variability in the renal clearance of many drug substrates. To date, a number of polymorphisms have been identified in the coding region and 5′ regulatory region of the *SLC22A6*, *SLC22A7*, *SLC22A8*, and *SLC22A11* gene variants, and some of them have caused changes in the transport activity of their encoded proteins. While different polymorphisms have been reported for genes encoding OAT—mainly SLC22A8 (encoding OAT3)—their allele frequency is $\leq 1\%$, and their functional significance has not been fully explained [21,39].

17.4 CONCLUSIONS

This chapter presented a broad outline of genetic variation affecting drug pharmacokinetics and consequently drug efficacy and toxicity that highlights the potential of pharmacogenomics to facilitate improved and more effective therapy. Important genetic associations that have been identified between variant genotypes and drug response phenotypes have prompted the FDA to revise labels for certain drugs to include relevant pharmacogenetic information and recommendations. For successful clinical implementation, there must be parallel development in genomic discovery and collaborative partnerships among stakeholders.

REFERENCES

[1] European Medicines Agency. Guideline on the use of pharmacogenetic methodologies in the pharmacokinetic evaluation of medicinal products. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/05/WC500090323.pdf>; 2010.

[2] US Food and Drug Administration. Table of valid genomic biomarkers in the context of approved drug labels. <http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm>; 2013.

[3] Williams JA, Andersson T, Andersson TB, Blanchard R, Behm MO, Cohen N, et al. PhRMA white paper on ADME pharmacogenomics. J Clin Pharmacol 2008;48(7):849–89.

[4] Ma Q, Lu AY. Pharmacogenetics, pharmacogenomics, and individualized medicine. Pharmacol Rev 2011;63(2):437–59.

[5] Weiss ST, McLeod HL, Flockhart DA, Dolan ME, Benowitz NL, Johnson JA, et al. Creating and evaluating genetic tests predictive of drug response. Nat Rev Drug Discov 2008;7(7):568–74.

[6] Karczewski KJ, Daneshjou R, Altman RB. Pharmacogenomics. PLoS Comput Biol 2012;8(12):e1002817.

[7] Pond SM, Tozer TN. First-pass elimination. Basic concepts and clinical consequences. Clin Pharmacokinet 1984;9(1):1–25.

[8] Conrado DJ, Gonzalez D, Derendorf H. Role of drug absorption in the pharmacokinetics of therapeutic interventions for stroke. Ann NY Acad Sci 2010;1207:134–42.

[9] Grossman I. ADME pharmacogenetics: current practices and future outlook. Expert Opin Drug Metab Toxicol 2009;5(5):449–62.

[10] Rescigno A. Area under the curve and bioavailability. Pharmacol Res 2000;42(6):539–40.

[11] Mondal D, Gerlach SL, Datta A, Chakravarty G, Abdel-Mageed AB. Pharmacogenomics dictate pharmacokinetics: polymorphisms in drug-metabolizing enzymes and drug-transporters. In: Noreddin Dr Ayman, editor. Readings in advanced pharmacokinetics—theory, methods and applications. InTech; 2012. p. 378. [20.4.12] ISBN 978-953-51-0533-6 [Chapter 4].

[12] Buxton ILO. Pharmacokinetics and pharmacodynamics: the dynamics of drug absorption, distribution, action, and elimination. In: Goodman & Gilman's the pharmacological basis of therapeutics. 11th ed. McGraw-Hill; 2006.

[13] Cipolla MJ. Control of cerebral blood flow. In: The cerebral circulation. Morgan & Claypool Life Sciences; 2009.<http://www.ncbi.nlm.nih.gov/books/NBK53082/>. Copyright © 2010, Bookshelf ID: NBK53082 [Chapter 5].

[14] Nayak-Rao S. Aminoglycoside use in renal failure. Indian J Nephrol 2010:121–4.

[15] Giacomini KM, Sugiyama Y. Membrane transporters and drug response. In: Goodman & Gilman's the pharmacological basis of therapeutics. 11th ed. McGraw-Hill; 2006.

[16] Colabufo NA, Berardi F, Contino M, Niso M, Perrone R. ABC pumps and their role in active drug transport. Curr Top Med Chem 2009;9(2):119–29.

[17] Mangoni AA, Jackson SH. Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications. Br J Clin Pharmacol 2004;57(1):6–14.

[18] Toutain PL, Bousquet-Melou A. Volumes of distribution. J Vet Pharmacol Ther 2004;27(6):441–53.

[19] Grover A, Benet LZ. Effects of drug transporters on volume of distribution. AAPS J 2009;11(2):250–61.

[20] Gonzalez FJ, Tukey RH. Drug metabolism. In: Brunton L, Lazo JS, Parker KL, editors. Goodman and Gilman's the pharmacological basis of therapeutics. 11th Ed. New York: McGraw-Hill; 2006. p. 82.

[21] Li J, Bluth MH. Pharmacogenomics of drug metabolizing enzymes and transporters: implications for cancer therapy. Pharmgenomics Pers Med. 2011;4:11–33.

[22] Trottier J, Verreault M, Grepper S, Monte D, Belanger J, Kaeding J, et al. Human UDP-glucuronosyltransferase (UGT)1A3 enzyme conjugates chenodeoxycholic acid in the liver. Hepatology 2006;44(5):1158–70.

[23] Ogu CC, Maxa JL. Drug interactions due to cytochrome P450. Proc (Bayl Univ Med Cent) 2000:421–3.

[24] Hande S, Wilson-Rich N, Bousvaros A, Zholudev A, Maurer R, Banks P, et al. 5-aminosalicylate therapy is associated with higher 6-thioguanine levels in adults and children with inflammatory bowel disease in remission on 6-mercaptopurine or azathioprine. Inflamm Bowel Dis 2006;12(4):251–7.

[25] Hayes JD, Flanagan JU, Jowsey IR. Glutathione transferases. Annu Rev Pharmacol Toxicol 2005;45:51–88.

[26] van Kuilenburg AB, Haasjes J, Richel DJ, Zoetekouw L, Van LH, De Abreu RA, et al. Clinical implications of dihydropyrimidine dehydrogenase (DPD) deficiency in patients with severe 5-fluorouracil-associated toxicity: identification of new mutations in the DPD gene. Clin Cancer Res 2000;6(12):4705–12.

[27] Gonzalez D, Conrado DJ, Theuretzbacher U, Derendorf H. The effect of critical illness on drug distribution. Curr Pharm Biotechnol 2011;12(12):2030–6.

[28] Park BK, Breckenridge AM. Clinical implications of enzyme induction and enzyme inhibition. Clin Pharmacokinet 1981;6(1):1–24.

[29] Johannessen SI, Landmark CJ. Antiepileptic drug interactions – principles and clinical implications. Curr Neuropharmacol 2010;8(3):254–67.

[30] Yap YG, Camm AJ. Drug induced QT prolongation and torsades de pointes. Heart 2003;89(11):1363–72.

[31] Brater DC. Measurement of renal function during drug development. Br J Clin Pharmacol 2002;54(1):87–95.

[32] Bendayan R. Renal drug transport: a review. Pharmacotherapy 1996;16(6):971–85.

[33] D'Angio RG, Gwilt PR. Time to steady state following a change in dosing rate. DICP 1989;23(6):468–72.

[34] Gardiner SJ, Begg EJ. Pharmacogenetics, drug-metabolizing enzymes, and clinical practice. Pharmacol Rev 2006;58(3):521–90.

[35] Zanger UM, Turpeinen M, Klein K, Schwab M. Functional pharmacogenetics/genomics of human cytochromes P450 involved in drug biotransformation. Anal Bioanal Chem 2008;392(6):1093–108.

[36] Belpaire FM, Bogaert MG. Cytochrome P450: genetic polymorphism and drug interactions. Acta Clin Belg 1996;51(4):254–60.

[37] Schneider E, Clark DS. Cytochrome P450 (CYP) enzymes and the development of CYP biosensors. Biosens Bioelectron 2013;39(1):1–13.

[38] Zhou SF, Liu JP, Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. Drug Metab Rev 2009;41(2):89–295.

[39] Lam YF, Cavallari LH. Principles of pharmacogenomics: pharmacokinetics, pharmacodynamics and clinical implications. Pharmacogenomics: Challenges and Opportunities in Therapeutic Implementation 2013:1–25. [chapter1].

[40] Bijl MJ, Visser LE, van Schaik RH, Kors JA, Witteman JC, Hofman A, et al. Genetic variation in the CYP2D6 gene is associated with a lower heart rate and blood pressure in beta-blocker users. Clin Pharmacol Ther 2009;85(1):45–50.

[41] Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. Science 1999;286(5439):487–91.

[42] Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. Pharmacogenomics J 2005;5(1):6–13.

[43] Cole PL, Beamer AD, McGowan N, Cantillon CO, Benfell K, Kelly RA, et al. Efficacy and safety of perhexiline maleate in refractory angina. A double-blind placebo-controlled clinical trial of a novel antianginal agent. Circulation 1990;81(4):1260–70.

[44] Hemeryck A, Lefebvre RA, De VC, Belpaire FM. Paroxetine affects metoprolol pharmacokinetics and pharmacodynamics in healthy volunteers. Clin Pharmacol Ther 2000;67(3):283–91.

[45] McGourty JC, Silas JH, Lennard MS, Tucker GT, Woods HF. Metoprolol metabolism and debrisoquine oxidation polymorphism–population and family studies. Br J Clin Pharmacol 1985;20(6):555–66.

[46] Fux R, Morike K, Prohmer AM, Delabar U, Schwab M, Schaeffeler E, et al. Impact of CYP2D6 genotype on adverse effects during treatment with metoprolol: a prospective clinical study. Clin Pharmacol Ther 2005;78(4):378–87.

[47] Terra SG, Pauly DF, Lee CR, Patterson JH, Adams KF, Schofield RS, et al. Beta-adrenergic receptor polymorphisms and responses during titration of metoprolol controlled release/extended release in heart failure. Clin Pharmacol Ther 2005;77(3):127–37.

[48] Zineh I, Beitelshees AL, Gaedigk A, Walker JR, Pauly DF, Eberst K, et al. Pharmacokinetics and CYP2D6 genotypes do not predict metoprolol adverse events or efficacy in hypertension. Clin Pharmacol Ther 2004;76(6):536–44.

[49] Santos PC, Dinardo CL, Schettert IT, Soares RA, Kawabata-Yoshihara L, Bensenor IM, et al. CYP2C9 and VKORC1 polymorphisms influence warfarin dose variability in patients on long-term anticoagulation. Eur J Clin Pharmacol 2013;69(4):789–97.

[50] Polasek TM, Lin FP, Miners JO, Doogue MP. Perpetrators of pharmacokinetic drug-drug interactions arising from altered cytochrome P450 activity: a criteria-based assessment. Br J Clin Pharmacol 2011;71(5):727–36.

[51] Lee CR, Goldstein JA, Pieper JA. Cytochrome P450 2C9 polymorphisms: a comprehensive review of the *in-vitro* and human data. Pharmacogenetics 2002;12(3):251–63.

[52] Kirchheiner J, Bauer S, Meineke I, Rohde W, Prang V, Meisel C, et al. Impact of CYP2C9 and CYP2C19 polymorphisms on tolbutamide kinetics and the insulin and glucose response in healthy volunteers. Pharmacogenetics 2002;12(2):101–9.

[53] Kirchheiner J, Brockmoller J, Meineke I, Bauer S, Rohde W, Meisel C, et al. Impact of CYP2C9 amino acid polymorphisms on glyburide kinetics and on the insulin and glucose response in healthy volunteers. Clin Pharmacol Ther 2002;71(4):286–96.

[54] Niemi M, Cascorbi I, Timm R, Kroemer HK, Neuvonen PJ, Kivisto KT. Glyburide and glimepiride pharmacokinetics in subjects with different CYP2C9 genotypes. Clin Pharmacol Ther 2002;72(3):326–32.

[55] Csajka C, Buclin T, Brunner HR, Biollaz J. Pharmacokinetic-pharmacodynamic profile of angiotensin II receptor antagonists. Clin Pharmacokinet 1997;32(1):1–29.

[56] Lee CR, Pieper JA, Hinderliter AL, Blaisdell JA, Goldstein JA. Losartan and E3174 pharmacokinetics in cytochrome P450 2C9*1/*1, *1/*2, and *1/*3 individuals. Pharmacotherapy 2003;23(6):720–5.

[57] Sekino K, Kubota T, Okada Y, Yamada Y, Yamamoto K, Horiuchi R, et al. Effect of the single CYP2C9*3 allele on pharmacokinetics and pharmacodynamics of losartan in healthy Japanese subjects. Eur J Clin Pharmacol 2003;59(8–9):589–92.

[58] Yasar U, Forslund-Bergengren C, Tybring G, Dorado P, Llerena A, Sjoqvist F, et al. Pharmacokinetics of losartan and its metabolite E-3174 in relation to the CYP2C9 genotype. Clin Pharmacol Ther 2002;71(1):89–98.

[59] Fischer TL, Pieper JA, Graff DW, Rodgers JE, Fischer JD, Parnell KJ, et al. Evaluation of potential losartan-phenytoin drug interactions in healthy volunteers. Clin Pharmacol Ther 2002;72(3):238–46.

[60] van der Weide J, Steijns LS, van Weelden MJ, de HK. The effect of genetic polymorphism of cytochrome P450 CYP2C9 on phenytoin dose requirement. Pharmacogenetics 2001;11(4):287–91.

[61] Aynacioglu AS, Brockmoller J, Bauer S, Sachse C, Guzelbey P, Ongen Z, et al. Frequency of cytochrome P450 CYP2C9 variants in a Turkish population and functional relevance for phenytoin. Br J Clin Pharmacol 1999;48(3):409–15.

[62] Lee AY, Kim MJ, Chey WY, Choi J, Kim BG. Genetic polymorphism of cytochrome P450 2C9 in diphenylhydantoin-induced cutaneous adverse drug reactions. Eur J Clin Pharmacol 2004;60(3):155–9.

[63] Sim SC, Risinger C, Dahl ML, Aklillu E, Christensen M, Bertilsson L, et al. A common novel CYP2C19 gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. Clin Pharmacol Ther 2006;79(1):103–13.

[64] Baldwin RM, Ohlsson S, Pedersen RS, Mwinyi J, Ingelman-Sundberg M, Eliasson E, et al. Increased omeprazole metabolism in carriers of the CYP2C19*17 allele; a pharmacokinetic study in healthy volunteers. Br J Clin Pharmacol 2008;65(5):767–74.

[65] Desta Z, Zhao X, Shin JG, Flockhart DA. Clinical significance of the cytochrome P450 2C19 genetic polymorphism. Clin Pharmacokinet 2002;41(12):913–58.

[66] Andersson T, Holmberg J, Rohss K, Walan A. Pharmacokinetics and effect on caffeine metabolism of the proton pump inhibitors, omeprazole, lansoprazole, and pantoprazole. Br J Clin Pharmacol 1998;45(4):369–75.

[67] Andersson T, Hassan-Alin M, Hasselgren G, Rohss K, Weidolf L. Pharmacokinetic studies with esomeprazole, the (S)-isomer of omeprazole. Clin Pharmacokinet 2001;40(6):411–26.

[68] Shirai N, Furuta T, Moriyama Y, Okochi H, Kobayashi K, Takashima M, et al. Effects of CYP2C19 genotypic differences in the metabolism of omeprazole and rabeprazole on intragastric pH. Aliment Pharmacol Ther 2001;15(12):1929–37.

[69] Furuta T, Ohashi K, Kosuge K, Zhao XJ, Takashima M, Kimura M, et al. CYP2C19 genotype status and effect of omeprazole on intragastric pH in humans. Clin Pharmacol Ther 1999;65(5):552–61.

[70] Furuta T, Shirai N, Xiao F, Ohashi K, Ishizaki T. Effect of high-dose lansoprazole on intragastic pH in subjects who are homozygous extensive metabolizers of cytochrome P4502C19. Clin Pharmacol Ther 2001;70(5):484–92.

[71] Sapone A, Vaira D, Trespidi S, Perna F, Gatta L, Tampieri A, et al. The clinical role of cytochrome p450 genotypes in *Helicobacter pylori* management. Am J Gastroenterol 2003;98(5):1010–5.

[72] Tanigawara Y, Aoyama N, Kita T, Shirakawa K, Komada F, Kasuga M, et al. CYP2C19 genotype-related efficacy of omeprazole for the treatment of infection caused by *Helicobacter pylori*. Clin Pharmacol Ther 1999;66(5):528–34.

[73] Sagar M, Janczewska I, Ljungdahl A, Bertilsson L, Seensalu R. Effect of CYP2C19 polymorphism on serum levels of vitamin B12 in patients on long-term omeprazole treatment. Aliment Pharmacol Ther 1999;13(4):453–8.

[74] Raunio H, Rautio A, Gullsten H, Pelkonen O. Polymorphisms of CYP2A6 and its practical consequences. Br J Clin Pharmacol 2001;52(4):357–63.

[75] Kubota T, Nakajima-Taniguchi C, Fukuda T, Funamoto M, Maeda M, Tange E, et al. CYP2A6 polymorphisms are associated with nicotine dependence and influence withdrawal symptoms in smoking cessation. Pharmacogenomics J 2006;6(2):115–9.

[76] Shimada T, Yamazaki H, Mimura M, Inui Y, Guengerich FP. Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. J Pharmacol Exp Ther 1994;270(1):414–23.

[77] Nolan D, Phillips E, Mallal S. Efavirenz and CYP2B6 polymorphism: implications for drug toxicity and resistance. Clin Infect Dis 2006;42(3):408–10.

[78] Turpeinen M, Tolonen A, Uusitalo J, Jalonen J, Pelkonen O, Laine K. Effect of clopidogrel and ticlopidine on cytochrome P450 2B6 activity as measured by bupropion hydroxylation. Clin Pharmacol Ther 2005;77(6):553–9.

[79] Martin H, Sarsat JP, de Waziers I, Housset C, Balladur P, Beaune P, et al. Induction of cytochrome P450 2B6 and 3A4 expression by phenobarbital and cyclophosphamide in cultured human liver slices. Pharm Res 2003;20(4):557–68.

[80] To KW, Liu ST, Cheung SW, Chan DP, Chan RC, Lee SS. Pharmacokinetics of plasma efavirenz and CYP2B6 polymorphism in southern Chinese. Ther Drug Monit 2009;31(4):527–30.

[81] Harris JW, Katki A, Anderson LW, Chmurny GN, Paukstelis JV, Collins JM. Isolation, structural determination, and biological activity of 6 alpha-hydroxytaxol, the principal human metabolite of taxol. J Med Chem 1994;37(5):706–9.

[82] Martinez C, Garcia-Martin E, Blanco G, Gamito FJ, Ladero JM, Agundez JA. The effect of the cytochrome P450 CYP2C8 polymorphism on the disposition of (R)-ibuprofen enantiomer in healthy subjects. Br J Clin Pharmacol 2005;59(1):62–9.

[83] Garcia-Martin E, Martinez C, Tabares B, Frias J, Agundez JA. Interindividual variability in ibuprofen pharmacokinetics is related to interaction of cytochrome P450 2C8 and 2C9 amino acid polymorphisms. Clin Pharmacol Ther 2004;76(2):119–27.

[84] Bahadur N, Leathart JB, Mutch E, Steimel-Crespi D, Dunn SA, Gilissen R, et al. CYP2C8 polymorphisms in Caucasians and their relationship with paclitaxel 6alpha-hydroxylase activity in human liver microsomes. Biochem Pharmacol 2002;64(11):1579–89.

[85] Niemi M, Backman JT, Neuvonen M, Neuvonen PJ. Effects of gemfibrozil, itraconazole, and their combination on the pharmacokinetics and pharmacodynamics of repaglinide: potentially hazardous interaction between gemfibrozil and repaglinide. Diabetologia 2003;46(3):347–51.

[86] Lopez-Rodriguez R, Novalbos J, Gallego-Sandin S, Roman-Martinez M, Torrado J, Gisbert JP, et al. Influence of CYP2C8 and CYP2C9 polymorphisms on pharmacokinetic and pharmacodynamic parameters of racemic and enantiomeric forms of ibuprofen in healthy volunteers. Pharmacol Res 2008;58(1):77–84.

[87] Lin YS, Lockwood GF, Graham MA, Brian WR, Loi CM, Dobrinska MR, et al. *In-vivo* phenotyping for CYP3A by a single-point determination of midazolam plasma concentration. Pharmacogenetics 2001;11(9):781–91.

[88] Hustert E, Haberl M, Burk O, Wolbold R, He YQ, Klein K, et al. The genetic determinants of the CYP3A5 polymorphism. Pharmacogenetics 2001;11(9):773–9.

[89] Xie HG, Wood AJ, Kim RB, Stein CM, Wilkinson GR. Genetic variability in CYP3A5 and its possible consequences. Pharmacogenomics 2004;5(3):243–72.

[90] Servedio V, d'Apolito M, Maiorano N, Minuti B, Torricelli F, Ronchi F, et al. Spectrum of UGT1A1 mutations in Crigler-Najjar (CN) syndrome patients: identification of twelve novel alleles and genotype-phenotype correlation. Hum Mutat 2005;25(3):325.

[91] Iyer L, King CD, Whitington PF, Green MD, Roy SK, Tephly TR, et al. Genetic predisposition to the metabolism of irinotecan (CPT-11). Role of uridine diphosphate glucuronosyltransferase isoform 1A1 in the glucuronidation of its active metabolite (SN-38) in human liver microsomes. J Clin Invest 1998;101(4):847–54.

[92] Bosma PJ, Chowdhury JR, Bakker C, Gantla S, de BA, Oostra BA, et al. The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. N Engl J Med 1995;333(18):1171–5.

[93] Ando Y, Saka H, Ando M, Sawa T, Muro K, Ueoka H, et al. Polymorphisms of UDP-glucuronosyltransferase gene and irinotecan toxicity: a pharmacogenetic analysis. Cancer Res 2000;60(24):6921–6.

[94] Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. Am J Hum Genet 1980;32(5):651–62.

[95] Schaeffeler E, Fischer C, Brockmeier D, Wernet D, Moerike K, Eichelbaum M, et al. Comprehensive analysis of thiopurine S-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel TPMT variants. Pharmacogenetics 2004;14(7):407–17.

[96] Colombel JF, Ferrari N, Debuysere H, Marteau P, Gendre JP, Bonaz B, et al. Genotypic analysis of thiopurine S-methyltransferase in patients with Crohn's disease and severe myelosuppression during azathioprine therapy. Gastroenterology 2000;118(6):1025–30.

[97] Priest VL, Begg EJ, Gardiner SJ, Frampton CM, Gearry RB, Barclay ML, et al. Pharmacoeconomic analyses of azathioprine, methotrexate and prospective pharmacogenetic testing for the management of inflammatory bowel disease. Pharmacoeconomics 2006;24(8):767–81.

[98] Lu Z, Zhang R, Diasio RB. Dihydropyrimidine dehydrogenase activity in human peripheral blood mononuclear cells and liver: population characteristics, newly identified deficient patients, and clinical implication in 5-fluorouracil chemotherapy. Cancer Res 1993;53(22):5433–8.

[99] Cascorbi I, Drakoulis N, Brockmoller J, Maurer A, Sperling K, Roots I. Arylamine N-acetyltransferase (NAT2) mutations and their

allelic linkage in unrelated Caucasian individuals: correlation with phenotypic activity. Am J Hum Genet 1995;57(3):581–92.

[100] Xie HG, Xu ZH, Ou-Yang DS, Shu Y, Yang DL, Wang JS, et al. Meta-analysis of phenotype and genotype of NAT2 deficiency in Chinese populations. Pharmacogenetics 1997;7(6):503–14.

[101] Borst P, Elferink RO. Mammalian ABC transporters in health and disease. Annu Rev Biochem 2002;71:537–92.

[102] Borst P, Evers R, Kool M, Wijnholds J. A family of drug transporters: the multidrug resistance-associated proteins. J Natl Cancer Inst 2000;92(16):1295–302.

[103] Lepper ER, Nooter K, Verweij J, Acharya MR, Figg WD, Sparreboom A. Mechanisms of resistance to anticancer drugs: the role of the polymorphic ABC transporters ABCB1 and ABCG2. Pharmacogenomics 2005;6(2):115–38.

[104] Sissung TM, Baum CE, Kirkland CT, Gao R, Gardner ER, Figg WD. Pharmacogenetics of membrane transporters: an update on current approaches. Mol Biotechnol 2010;44(2):152–67.

[105] Cascorbi I, Haenisch S. Pharmacogenetics of ATP-binding cassette transporters and clinical implications. Methods Mol Biol 2010;596:95–121.

[106] Cascorbi I. P-glycoprotein: tissue distribution, substrates, and functional consequences of genetic variations. Handb Exp Pharmacol 2011;201:261–83.

[107] Lai Y, Varma M, Feng B, Stephens JC, Kimoto E, El-Kattan A, et al. Impact of drug transporter pharmacogenomics on pharmacokinetic and pharmacodynamic variability – considerations for drug development. Expert Opin Drug Metab Toxicol 2012;8(6):723–43.

[108] Sakaeda T. MDR1 genotype-related pharmacokinetics: fact or fiction? Drug Metab Pharmacokinet 2005;20(6):391–414.

[109] Dulucq S, Bouchet S, Turcq B, Lippert E, Etienne G, Reiffers J, et al. Multidrug resistance gene (MDR1) polymorphisms are associated with major molecular responses to standard-dose imatinib in chronic myeloid leukemia. Blood 2008;112(5):2024–7.

[110] Yang Z, Wu D, Bui T, Ho RJ. A novel human multidrug resistance gene MDR1 variant G571A (G191R) modulates cancer drug resistance and efflux transport. J Pharmacol Exp Ther 2008;327(2):474–81.

[111] Franke RM, Gardner ER, Sparreboom A. Pharmacogenetics of drug transporters. Curr Pharm Des 2010;16(2):220–30.

[112] Choi YH, Yu AM. ABC transporters in multidrug resistance and pharmacokinetics, and strategies for drug development. Curr Pharm Des 2013. [Epub ahead of print]

[113] Naesens M, Kuypers DR, Verbeke K, Vanrenterghem Y. Multidrug resistance protein 2 genetic polymorphisms influence mycophenolic acid exposure in renal allograft recipients. Transplantation 2006;82(8):1074–84.

[114] Rau T, Erney B, Gores R, Eschenhagen T, Beck J, Langer T. High-dose methotrexate in pediatric acute lymphoblastic leukemia: impact of ABCC2 polymorphisms on plasma concentrations. Clin Pharmacol Ther 2006;80(5):468–76.

[115] Han JY, Lim HS, Yoo YK, Shin ES, Park YH, Lee SY, et al. Associations of ABCB1, ABCC2, and ABCG2 polymorphisms with irinotecan-pharmacokinetics and clinical outcome in patients with advanced non-small cell lung cancer. Cancer 2007;110(1):138–47.

[116] Haenisch S, May K, Wegner D, Caliebe A, Cascorbi I, Siegmund W. Influence of genetic polymorphisms on intestinal expression and rifampicin-type induction of ABCC2 and on bioavailability of talinolol. Pharmacogenet Genomics 2008;18(4):357–65.

[117] Izzedine H, Hulot JS, Villard E, Goyenvalle C, Dominguez S, Ghosn J, et al. Association between ABCC2 gene haplotypes and tenofovir-induced proximal tubulopathy. J Infect Dis 2006;194(11):1481–91.

[118] Niemi M, Arnold KA, Backman JT, Pasanen MK, Godtel-Armbrust U, Wojnowski L, et al. Association of genetic polymorphism in ABCC2 with hepatic multidrug resistance-associated protein 2 expression and pravastatin pharmacokinetics. Pharmacogenet Genomics 2006;16(11):801–8.

[119] Kusuhara H, Sugiyama Y. ATP-binding cassette, subfamily G (ABCG family). Pflugers Arch 2007;453(5):735–44.

[120] Giacomini KM, Balimane PV, Cho SK, Eadon M, Edeki T, Hillgren KM, et al. International transporter consortium commentary on clinically important transporter polymorphisms. Clin Pharmacol Ther 2013;94(1):23–6.

[121] Tamai I, Nezu J, Uchino H, Sai Y, Oku A, Shimane M, et al. Molecular identification and characterization of novel members of the human organic anion transporter (OATP) family. Biochem Biophys Res Commun 2000;273(1):251–60.

[122] Nakanishi T, Tamai I. Genetic polymorphisms of OATP transporters and their impact on intestinal absorption and hepatic disposition of drugs. Drug Metab Pharmacokinet 2012;27(1):106–21.

[123] Tamai I. Oral drug delivery utilizing intestinal OATP transporters. Adv Drug Deliv Rev 2012;64(6):508–14.

[124] Shitara Y, Maeda K, Ikejiri K, Yoshida K, Horie T, Sugiyama Y. Clinical significance of organic anion transporting polypeptides (OATPs) in drug disposition: their roles in hepatic clearance and intestinal absorption. Biopharm Drug Dispos 2013;34(1):45–78.

[125] Pasanen MK, Fredrikson H, Neuvonen PJ, Niemi M. Different effects of SLCO1B1 polymorphism on the pharmacokinetics of atorvastatin and rosuvastatin. Clin Pharmacol Ther 2007;82(6):726–33.

[126] Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy–a genomewide study. N Engl J Med 2008;359(8):789–99.

[127] Kullak-Ublick GA, Ismair MG, Stieger B, Landmann L, Huber R, Pizzagalli F, et al. Organic anion-transporting polypeptide B (OATP-B) and its functional comparison with three other OATPs of human liver. Gastroenterology 2001;120(2):525–33.

[128] Kimura N, Masuda S, Tanihara Y, Ueo H, Okuda M, Katsura T, et al. Metformin is a superior substrate for renal organic cation transporter OCT2 rather than hepatic OCT1. Drug Metab Pharmacokinet 2005;20(5):379–86.

[129] Grzybowska M, Bober J, Olszewska M. Metformin – mechanisms of action and use for the treatment of type 2 diabetes mellitus. Postepy Hig Med Dosw (Online) 2011;65:277–85.

[130] Filipski KK, Mathijssen RH, Mikkelsen TS, Schinkel AH, Sparreboom A. Contribution of organic cation transporter 2 (OCT2) to cisplatin-induced nephrotoxicity. Clin Pharmacol Ther 2009;86(4):396–402.

Chapter 18

Pharmacodynamic Pharmacogenomics

Linsay McCallum, Stefanie Lip, and Sandosh Padmanabhan
BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

Chapter Outline

18.1 Overview 365
18.2 Basic Pharmacodynamics 365
  18.2.1 Historical Aspects 365
  18.2.2 Basic Principles 366
  18.2.3 Receptors 366
  18.2.4 Measurement of Drug Action 370
18.3 Pharmacodynamic Pharmacogenetics 373
  18.3.1 Pharmacodynamics and Targets for Drug Action 373
  18.3.2 Genetic Variability that Indirectly Alters Drug Response 376
  18.3.3 Adverse Drug Reactions 378
18.4 Conclusion: Pharmacodynamic Pharmacogenomics into Clinical Practice 379
Glossary 380
Acronyms and Abbreviations 381
References 381

18.1 OVERVIEW

Drug response or adverse effect is the net effect of multiple factors: age, organ function, concomitant therapy, drug interactions, and disease. In addition to these, there are now numerous examples of cases where interindividual differences in drug response are due to sequence variants in genes encoding drug-metabolizing enzymes, drug transporters, or drug targets. The study of the biochemical and physiological interactions between drugs and their targets and the relationship between plasma drug concentration and drug effect is known as pharmacodynamics. The genetic determinants of drug response are potentially of great clinical value as, unlike other factors influencing drug response, they generally remain stable throughout a person's lifetime. This chapter focuses on the principles of pharmacodynamics and how genetic variation of drug targets can have an effect on drug response.

18.2 BASIC PHARMACODYNAMICS

Pharmacodynamics describes the relationship between the drug concentration at the receptor and the physiological and biochemical effect produced. Pharmacodynamics is often described as "what the drug does to the body."

18.2.1 Historical Aspects

> *If you want to explain any poison properly, what then isn't a poison? All things are poison, nothing is without poison; the dose alone causes a thing not to be a poison.*
>
> Paracelsus 1493–1541 A.D.

Medicines have been used for thousands of years, initially in the form of minerals and plants observed to have therapeutic effects without any understanding of the mechanism behind them. Paracelsus (Theophrastus von Hohenheim, 1493–1541) was a Swiss physician, botanist, alchemist, and astrologer. He was the first to question the effects of prescribed medicines and sought knowledge of the active ingredients and the theory behind their mechanisms of action while rejecting irrational compounds. Paracelsus went on to explain the initial effects of dose and adverse effects, and this was the beginning of what has grown to become the modern pharmacodynamic theories of dose-concentration-response relationships and the concept of the therapeutic index.

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00018-9
© 2014 Elsevier Inc. All rights reserved.

Johann Jakob Wepfer (1620–1695), a Swiss pharmacologist and pathologist, was the first to use animal experimentation to verify theories about pharmacological actions, studying the effects of water hemlock, hellebore, and monkshood. His studies paved the way for the current practice of preclinical studies to test new drugs in humans.

Human studies were carried out in 1805 by Friedrich Sertuner, a German pharmacist who isolated the first alkaloid from opium and administered a large dose to himself and three acquaintances, all of whom suffered symptoms of severe opioid toxicity.

Oswald Schmeideberg (1838–1921), a Russian pharmacologist, helped to establish pharmacology as an independent science discipline. His thesis concerned the measurement of chloroform in blood, and he earned a medical doctorate at the University of Dorpat. Together with his working group, he discovered and described fundamental concepts in clinical pharmacology and pharmacodynamics, including structure–activity relationships (T. Frazer 1841–1921), drug–receptor relationships (J. Langley 1852–1925), and the emerging field of toxicology (P. Ehrlich 1854–1915). Alexander J Clark, an English pharmacologist, formalized receptor theory when the law of mass action was applied to drug–receptor theory in the 1920s.

The organ bath is currently used as standard for *in vitro* pharmacodynamic dose–response experiments involving the suspension of an isolated section of organ/tissue in a controlled solution with sensors to measure the contraction or relaxation of the tissue when drugs are added. The organ bath experiment was used by Henrick Magnus (1802–1870), who applied the method to a strip of small intestine. He was followed by Jean Francois Heymans, who in 1904 worked with the mammalian heart.

The first chair in pharmacology was established in 1890 in the United States at the University of Michigan under John Jacob Abel, who was an American pharmacologist trained by Oswald Schmiedeberg. Since the 1960s, departments of pharmacology and clinical pharmacology have been established in many universities.

As knowledge of clinical chemistry and pharmacology grew, the principles of these disciplines were applied to the pharmaceutical industry and particularly in drug development. Drugs are now designed to interact with known targets to produce therapeutic effects. Squibb developed captopril, whose structure was designed to fit the active site of the angiotensin-converting enzyme. This lead to the further development of the current ACE inhibitor class of drugs.

Pharmacology and the study of pharmacodynamics principles have been prominent throughout history, and the evolution of pharmacogenomics provides a new and exciting field.

18.2.2 Basic Principles

The functional macromolecular component of an organism that is presumed to interact with the drug is termed a receptor. However, even if a drug interacts with a specific molecular receptor and stable drug and metabolite concentrations are maintained, it does not imply that the drug effect will be constant over time. This is because the drug–receptor interaction is influenced by multiple factors, including physiologic milieu, concomitant disease, and drug effect. For example, receptors may be up- or downregulated by either the disease or the drug itself affecting drug efficacy and toxicity.

18.2.3 Receptors

Receptors are the key part of the signaling system that enable coordination of the activities of cells and organs. A receptor is a macromolecule, or an assembly of macromolecules, located in the cell membrane or within the cell that binds ligands (endogenous molecules or exogenous drugs) resulting in chemical signaling between and within cells with consequent pharmacological effects [1]. There are 646 genes, to date, that are known to encode for receptors [2]. The International Union of Basic and Clinical Pharmacology (IUPHAR) database contains pharmacological, physiological, and genetic information on human and rodent receptors, other drug targets, and the substances that act on them [2]. It is updated regularly.

Physiological receptors are proteins that act as receptors for endogenous ligands, and they tend to be highly selective. Drugs that act on physiological receptors can mimic the response of the endogenous ligand (agonist) or bind to the receptor, causing no effect but preventing endogenous ligand binding (antagonist), or they can mimic the response of the endogenous ligand to a limited extent but also prevent any further stimulation (partial agonist) [1]. Figures 18.1 through 18.3 and Table 18.1 illustrate the effect of agonists and antagonists on receptors. The strength of the reversible interaction between the drug and the receptor (the agonist dissociation constant (K_D)) is defined as the affinity of one for another [1]. Affinity is determined by the drug's chemical structure.

18.2.3.1 Mechanisms by Which Drugs Act on Receptors

Drugs can interfere with the usual function of the receptor at a site determined by the receptor's location. The effect of the drug is partly determined by the functional capacity of the receptor [1]. Receptors can mediate functions common to many cells and consequently the effect of a drug on a receptor is widespread and may result in more frequent unwanted effects [1]. If a drug acts on a receptor that is unique to one or a few cells, the effects will be more specific; however, the physiological consequences may still be widespread [1]. The ideal drug acts only in the receptor of interest at all concentrations and does not have any interaction with other receptors or biological systems at any concentration. However, there are no drugs with this ideal property.

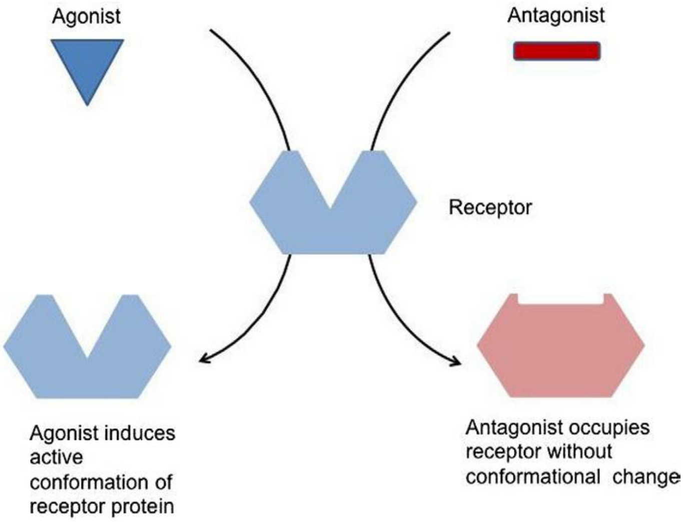

Agonist
Antagonist
Receptor
Agonist induces active conformation of receptor protein
Antagonist occupies receptor without conformational change

FIGURE 18.1 Action of agonist and antagonist on receptor.

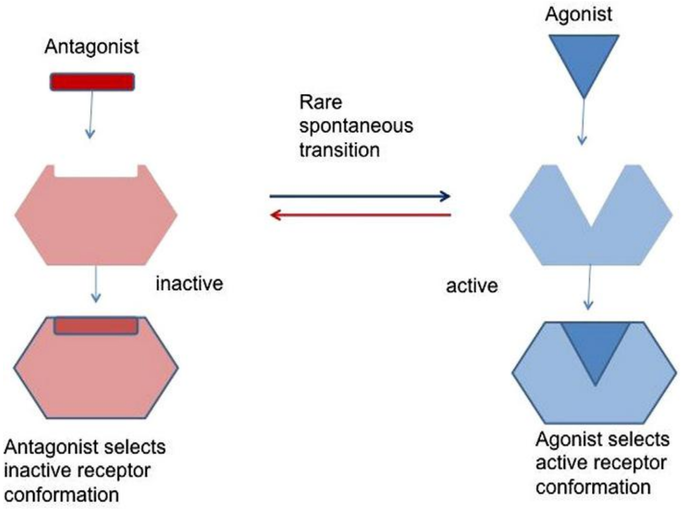

Antagonist
Agonist
Rare spontaneous transition
inactive
active
Antagonist selects inactive receptor conformation
Agonist selects active receptor conformation

FIGURE 18.2 Conformation of receptor following agonist and antagonist binding.

Receptors have two functions with specific domains: ligand binding (ligand-binding domain) and cell signaling (effector domain) [1]. The actions of a receptor may be directly on cellular targets, effector proteins, or intermediary cellular signal molecules (transducers) [1]. The receptor, target, and intermediary signaling molecule together are referred to as the signal transduction pathway. Often, the effector protein acts as an enzyme or transport protein that creates, moves, or breaks down a metabolite or ion. In this case, it is known as a second messenger. A second messenger is a molecule that relays signals from receptors to target molecules within the cell.

Second Messengers

An agonist binding to a receptor, considered the first messenger in receptor signal transduction, promotes mobilization of a second messenger [1]. Second messengers are molecules that relay signals from receptors on the cell surface to target molecules in the cytoplasm or nucleus [1]. They amplify the strength of the signal and can influence each other's metabolism [1]. Second messengers' essential properties are summarized as follows:

- The increase in second messenger levels is produced by the appropriate agonists and blocked by specific antagonists.
- The increase in second messenger levels precedes the increase in response.
- The response to an agonist can be mimicked by an exogenously applied second messenger.
- Substances that block the breakdown of the second messenger potentiate the response.

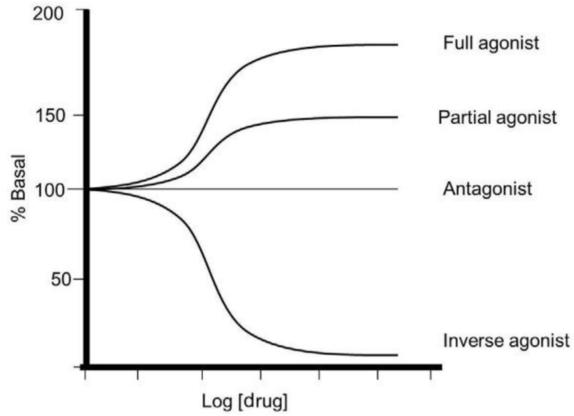


FIGURE 18.3 Action of full agonist, partial agonist, antagonist, and reverse agonist.

18.2.3.2 Types of Receptors

Receptors are classified into four types that exhibit characteristic properties. These are described in the following subsections.

Receptor Protein Kinases and Guanylyl Cyclases

Cell surface protein kinases are a group of receptors with intrinsic enzymatic activity. They exert their effect by phosphorylating diverse effector proteins on the inner side of the plasma membrane [1]. Phosphorylation is a common mechanism for facilitating the biochemical activities of an effector or its interaction with other proteins [1]. The majority of receptor protein kinases are composed of an agonist binding site on the extracellular surface of the plasma membrane, a solitary membrane-spanning element, and a protein kinase domain on the inner surface of the membrane. Other variations include the assembly of multiple subunits or multiple binding domains. Figure 18.4 illustrates the different receptor types and structures.

TABLE 18.1 Agonist and Antagonist Effects

Action	Effect
Agonist	Stimulates or activates the receptor to produce an effect
Antagonist	Prevents the action of an agonist but does not have any effect itself
Partial agonist	Stimulates the receptor to a limited extent but also prevents any further stimulation by naturally occurring agonists
Inverse agonist	Produces an effect opposite to that of the agonist

Receptors that bind guanylin and uroguanylin have a guanylyl cyclase intracellular domain, rather than a protein kinase domain, that synthesizes the second messenger molecule cyclic guanosine monophosphate (cGMP). This activates a cGMP-dependent protein kinase and can modulate the effects of effectors.

G protein Coupled Receptors

G protein–coupled receptors (GPCRs) are the largest family of receptors. They have distinct GTP-binding regulatory proteins known as G proteins. These act as cell-signaling molecules conveying information from the GPCR to a single or multiple effector proteins [1]. GPCRs are the target for many drugs and span the plasma membrane with seven transmembrane domains, an extracellular N-terminal domain, and an intracellular C-terminal domain [1]. The structure and function of GPCRs are shown in Figure 18.4. The G protein families and their effects are listed in Table 18.3; the three GPCR families are summarized in Table 18.4. The targets of GPCRs are summarized in Table 18.5.

For example, catecholamines (first messengers) activate β adrenoceptors to increase, through a G-protein coupled system, the activity of intracellular adenylate cyclase, which increases the rate of formation of cAMP (second messenger), which is a modulator of the activity of several enzyme systems that cause the cell to act.

Ion Channels

Ion channels allow the selective passage of selected ions through the cell membrane and can be opened or closed through a number of different mechanisms [1]. The receptor can exist in an open or closed state. Drugs may open or close the ion channel by binding to the channel protein, or they may affect the function indirectly by involving a G protein or another intermediary [1]. There are two main types of ion channel: ligand-gated ion channels, which open when an agonist is bound and can be considered a receptor; and voltage-gated ion channels, which open in response to alterations in transmembrane potential [1]. Voltage-gated ion channels are not strictly receptors as they do not require agonist binding to open and can be considered drug targets [1].

Ion channels are large structures with several similar or identical domains organized in an array of subunits or as a single large protein. Each subunit contains 2–6 transmembrane domains. Ion channels are selective to particular ion species, generally cations (Na^+, Ca^{2+}, K^+) or anions (Cl^-). They are the only way by which ions can cross the cell membrane, as they are unable to penetrate the lipid bilayer [1]. Ion channels are made up of protein molecules that form water-filled pores that span the cell membrane and can be opened or closed [1]. Different ions pass through different ion channels at different speeds depending on the electrochemical gradient. The gating and permeation of the ion channels is modulated by ligands binding to various sites

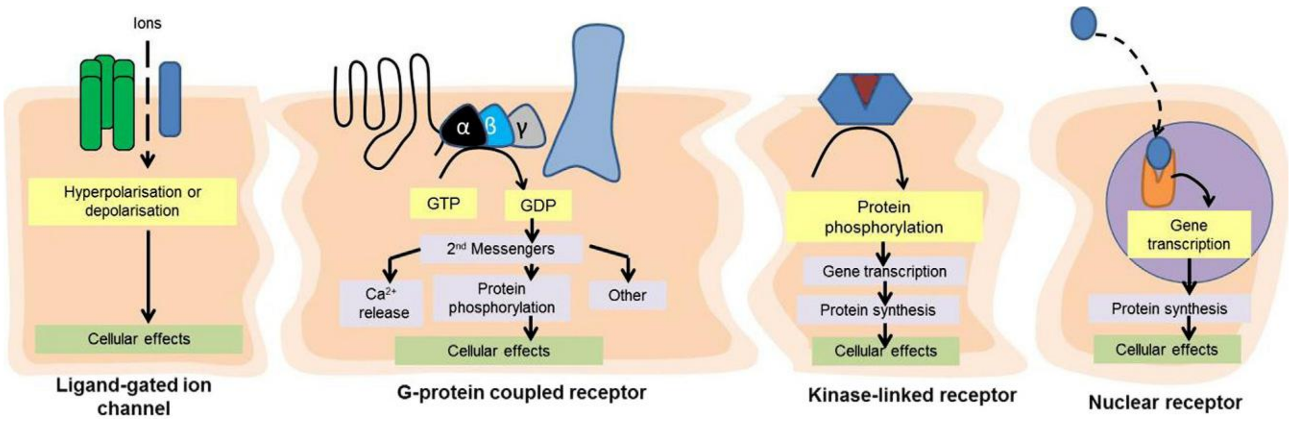

Ions
Hyperpolarisation or depolarisation
Cellular effects
Ligand-gated ion channel
α
β
γ
GTP
GDP
2nd Messengers
Ca2+ release
Protein phosphorylation
Other
Cellular effects
G-protein coupled receptor
Protein phosphorylation
Gene transcription
Protein synthesis
Cellular effects
Kinase-linked receptor
Gene transcription
Protein synthesis
Cellular effects
Nuclear receptor

FIGURE 18.4 Receptor types and their structure and function.

on the channel protein, mediators, and drugs acting in an indirect way, often through GPCR and second messengers in intracellular signaling pathways [1]. The structure and effects of a ligand-gated ion channel is shown in Figure 18.4.

Nuclear Receptors

Nuclear receptors are monomeric proteins with a core domain, N-terminal domain, C-terminal domain, and activation function site [1]. There are two classes of nuclear receptor (see Table 18.2). Steroid (estrogen, glucocorticoids) receptors located in the cell cytoplasm are translocated into the nucleus following binding with their steroid partner or located within the nucleus [1].

TABLE 18.2 Classes of Nuclear Receptor

Class I	Class II
Present in cytoplasm	Present in nucleus
Operate as homodynes	Operate as heterodimers
Endocrine ligands	Lipid ligands
High affinity	Low affinity
Glucocorticoid receptor, estrogen receptor	PPAR, RXR

TABLE 18.3 G Protein Families and their Effects

G protein	Effect
Gs	Activates adenylate cyclase
Gi/Go	Inhibits adenylate cyclase
Gq	Activates phospholipase C

Hormone response elements are short sequences of DNA to which nuclear receptors bind to modify gene transcription. Once inside the nucleus, the ligand-bound receptor modifies gene expression through its AF domains by recruiting co-activators. This is illustrated in Figure 18.4.

Enzymes

Interaction between a drug and an enzyme can be considered similar to that between a drug and a receptor. Drugs can alter enzyme activity because they are manufactured to resemble a natural substrate and therefore compete with the natural substrate for the enzyme-binding site. The action can be competitive and reversible in the case of enalapril. Enalapril is structurally similar to angiotensin I and occupies the site of angiotensin I on the angiotensin-converting enzyme (ACE), it thus inhibits ACE. Alternatively, the action can be irreversible in the case of aspirin, which combines covalently to cyclooxygenase (COX) and inhibits the enzyme in platelets for the duration of their lifespan [1].

Transport Receptors

Transport receptors are a specialized class of receptors transmitting molecular material rather than information. Transport of ions and small molecules across the cell membrane requires a carrier protein, as these molecules have insufficient lipid permeability or too high a polarity to cross the membrane by themselves [1]. The transport receptors have a specific recognition site that allows only a specific molecule or family of molecules to be transported. An example of a transport receptor is LDL.

18.2.3.3 Receptor Binding, Dose-Response, and Regulation

Receptor binding can be weak and rapidly reversible (e.g., hydrogen bonds) or strong and effectively irreversible (e.g., covalent bonds). Receptor numbers do not remain constant

TABLE 18.4 G Protein Coupled Receptor Families

Rhodopsin Family	Secretin/Glucagon Receptor Family	Glutamate Receptor/Calcium Sensor Family
Receptor for amine neurotransmitters, peptide, purines, prostanoids	Receptor for peptide hormones, secretin, glucagon	$GABA_B$ receptors, Ca^{2+}–sensing receptors
Short EC tail, ligand binds to transmembrane domain or EC loops	Intermediate EC tail incorporating ligand-binding domain	Long EC tail incorporating ligand-binding domain

but differ according to circumstances. When tissues are continuously or repeatedly exposed to an agonist, the number of receptors decreases (downregulation), which may be a cause of tachyphylaxis (the loss of efficacy with frequently repeated doses) [1]. The continuous or prolonged exposure to an antagonist leads to an increase in receptor numbers (upregulation), which is the reason for side effects associated with abrupt withdrawal of some drugs (e.g., β adrenoceptor blockers) [1].

The binding of drugs to receptors can often be measured directly by incubating samples of tissue with various concentrations of drugs, which are radioactively labeled (3H, 14C, 125I), and then waiting until equilibrium is reached. The amount of radioactivity is measured following the removal of the supernatant. There is usually some nonspecific binding where the radiolabeled drugs are taken up by structures other than the receptor of interest. The amount of nonspecific binding is estimated by a further experiment in which the radiolabeled drug is taken up in the presence of a saturating concentration of a nonradiolabeled receptor antagonist; the nonspecific binging proportion is measured. The specific binding is the total binding minus the nonspecific component. A binding curve is drawn to enable the affinity of the drug for the specified receptor and the binding capacity (B_{max}) to be estimated. This is illustrated in Figure 18.5 [1]. Binding curves for agonists often show drugs binding to two or more receptors with different affinities. Binding capacity is the density of receptors within the tissue. Positron emission tomography (PET) can be used to investigate the distribution of receptors within organs.

18.2.4 Measurement of Drug Action

The effect of a drug can be measured either by laboratory measurements or clinical outcomes. This section describes the relationship between drug dose and pharmacological response and mechanisms of drug response.

18.2.4.1 Dose–Response Relationship

The effect of a drug (e.g., drop in blood pressure or contraction of smooth muscle) can be determined using clinical or laboratory measurements. The results are then plotted against the dose or concentration that the recipient or section of tissue is exposed to. This is known as the dose–response curve *in vivo* and the concentration-effect curve *in vitro* [1]. A steep curve suggests that a

TABLE 18.5 Targets for G proteins

G Protein Target	Function
Adenylyl cyclase	cAMP formation
Phospholipase C	Inositol phosphate and DAG formation
Ion channels	Cation channels
Rho A/Rho kinase	System control of multiple signaling pathways that govern cell growth and proliferation

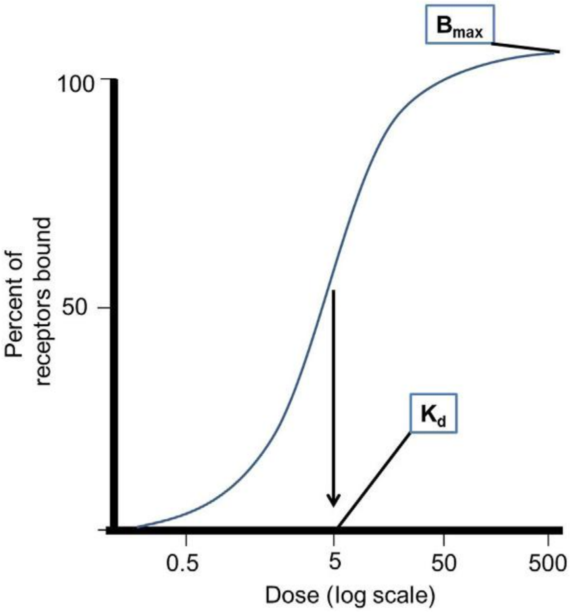


FIGURE 18.5 Binding curve showing Bmax.

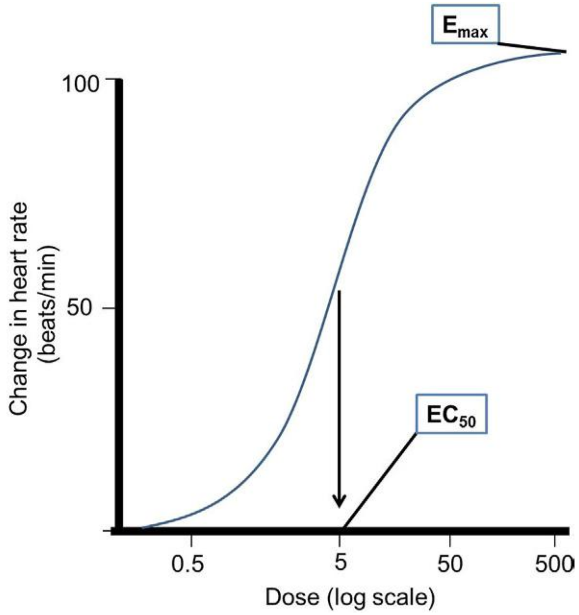


FIGURE 18.6 Dose–response curve showing ED_{50}.

small change in dose will produce large effects in clinical response or adverse effects; a flat relationship suggests that increasing the dose will provide little clinical advantage (see Figure 18.6). The dose-response curve and the concentration-effect curve allow estimation of the maximum response that the drug can produce (i.e., no further clinical effect will be gained by further increasing the dose) and the dose or concentration required to produce 50% of the maximum effect (ED_{50} or EC_{50}) [1]. Dose-response and concentration-effect curves cannot be used to estimate binding or affinity.

Although the concentration-effect curve is useful, it must be recognized that the concentration of the drug in solution may differ from its concentration at the receptors, as it will be subject to degradation, transport, and metabolism and may have already reached steady state.

18.2.4.2 Therapeutic Index

The therapeutic index is the relationship between the therapeutic and toxic dose of a given drug and is calculated using the ED_{50} and TD_{50}:

$$\text{Therapeutic Index} = TD_{50} / ED_{50}$$

ED_{50} is the dose required to produce a therapeutic effect in 50% of the population; TD_{50} is the dose required to produce a toxic effect in 50% of the population; both are calculated from dose–response curves. The therapeutic index is shown in Figure 18.7.

The dose of a drug can be increased progressively until a desired response is achieved; however if it is further increased, no additional desired effects are achieved and unwanted effects can be seen. Drugs with a narrow therapeutic index can be difficult to use in clinical practice and often require monitoring of plasma concentration in order to prevent toxicity [1]. Some examples of drugs with a narrow therapeutic index are as follows:

- Warfarin
- Lithium
- Digoxin
- Phenytoin
- Gentamicin
- Amphoteracin B
- 5-fluouracil
- Zidovudine

18.2.4.3 Competitive Antagonism

Unlike in clinical trials, the reality of modern medicine is that a drug's effect cannot be considered in isolation, as a large number of patients will be exposed to multiple drugs that can influence each other's effect in multiple ways. One common mode of drug interaction is competition for binding at the receptor site. When a competitive antagonist is

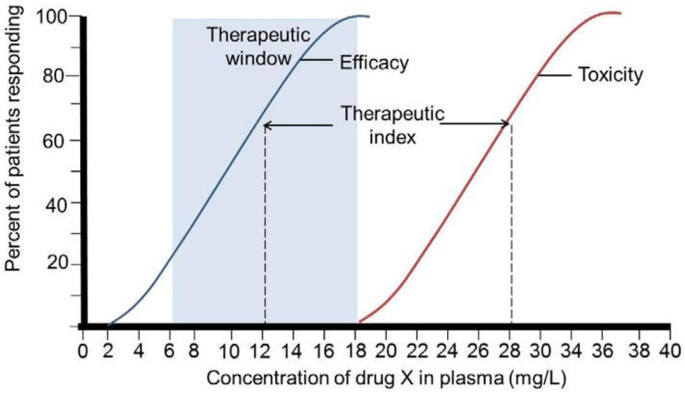


FIGURE 18.7 Therapeutic index.

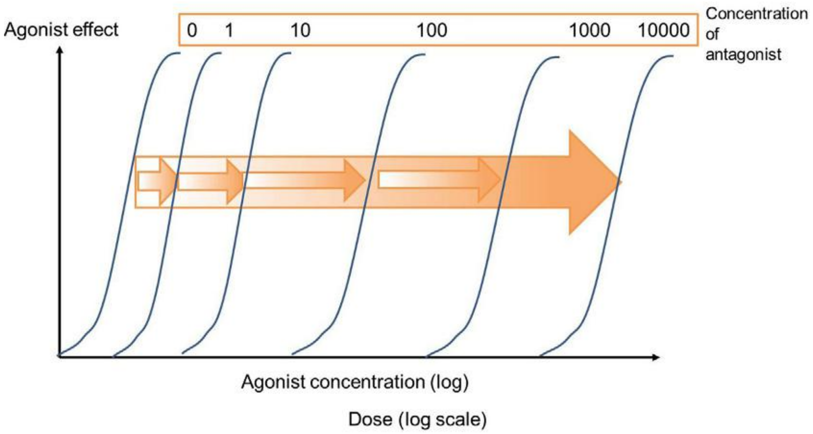


FIGURE 18.8 Agonist levels versus agonist effects in the presence of increasing antagonist levels.

present, the concentration-effect curve is shifted to the right with no change in slope or maximum concentration; this principle is illustrated in Figure 18.8 [1].

The agonist occupancy rate at a given agonist concentration, in the presence of a competitive antagonist, will be reduced because the receptor can bind only one molecule at a time. Given the nature of competition, increasing the antagonist concentration can restore agonist occupancy, making the competitive antagonism surmountable. In noncompetitive antagonism, increasing the agonist concentration does not restore agonist occupancy, making the antagonist effect insurmountable. Displacement of the competitive antagonist occurs because the agonist, by occupying the vacant receptors, reduces the rate of antagonist binding. Irreversible antagonism does occur (e.g., aspirin and monoamine oxidase inhibitors), and these drugs bond covalently to the receptor. The characteristics of competitive antagonism are as follows:

- Evidence of competition displayed during binding studies
- Shift of the concentration-effect curve to the right, without change in slope or maximum effect
- Linear relationship between agonist dose ratio and antagonist concentration

18.2.4.4 *Allosteric Modulation*

In addition to the specified agonist binding site, receptors can also have many other binding sites through which drugs can alter receptor function by increasing or decreasing receptor affinity for agonist binding or by altering efficacy [2]. These binding sites are referred to as allosteric, and ligands may be allosteric facilitators of the agonist effect or allosteric antagonists [2].

18.2.4.5 *Efficacy and Potency*

Efficacy is related to the ability of the agonist to activate the receptor. Full agonists have a high efficacy while antagonists have a low efficacy. Efficacy is related to receptor occupancy: compounds with low efficacy are required to occupy more receptors to produce the same response compared to those with higher efficacy. Full agonist efficacy is variable, as different agonists may require occupation of a different number of receptors to produce a maximum response. Intrinsic activity refers to the relative ability of a drug receptor complex to produce a maximum response effect. The intrinsic activity of a full agonist is always 1 while the intrinsic activity of an antagonist is always 0 because they bind to receptors but do not produce any effect. Partial agonists have an intrinsic activity between 0 and 1.

Potency is a measure of drug activity expressed in terms of the quantity required to produce a given effect—that is, the amount (weight) of a drug in relation to its effect. For example, if the weight for drug X has a greater effect than the weight of drug Y, even if the maximum therapeutic effect is similar with both drugs, drug X is said to be more potent [1].

18.2.4.6 *Desensitization and Tachyphylaxis*

Repeated exposure to a drug may result in a reduction of effect. Receptor numbers do not remain constant but differ according to circumstances. This can be due to receptor up-or downregulation. Desensitization and tachyphylaxis can occur very quickly; however, tolerance describes a reduced effect over a longer period of time [1]. The mechanisms by which desensitization, tachyphylaxis, and tolerance occur are described in the following sections.

Change in Receptors

Desensitization occurs rapidly in receptors coupled to ion channels and is caused by a conformational change in the receptor that results in tight binding of the agonist without opening of the channel [1]. A second mechanism is the phosphorylation of intracellular regions of the receptor protein, which is a slower mechanism. GPCRs also show desensitization. Second messenger cascades are not activated, although the agonist is bound, because receptor phosphorylation occurs [1]. This uncoupling mechanism happens over a few minutes.

Translocation of Receptors

Repeated or prolonged exposure to a drug results in a gradual decrease in receptor numbers expressed on the cell surface. This is a result of internalization of the receptors and occurs within hours. The internalized receptors are moved into the cell by endocytosis of sections of the cell membrane. Internalization is phosphorylation dependent [1].

Alterations in Drug Metabolism

Tolerance to some drugs occurs because repeated exposure to the same dose causes a lowering of plasma concentrations as a result of metabolic degradation [1].

Exhaustion of Mediators and Physiological Adaptation

If essential intermediate substances are depleted, desensitization can occur. For example, a drug's effect on the body may be nullified by homeostatic responses, causing desensitization [1].

18.2.4.7 Non-Receptor-Mediated Drug Response

Although the majority of drugs act through receptors, some exert their effects through other pathways or simply by being in an appropriate area of the body. Osmotic laxatives draw fluid from the bowel without binding. Chelating agents bind directly with heavy metals. Osmotic diuretics (mannitol) inhibit reabsorption of Na^{2+} and water in the renal proximal tubule by increasing osmolality of the tubule fluid, resulting in volume expansion.

18.3 PHARMACODYNAMIC PHARMACOGENETICS

Pharmacodynamic pharmacogenomics is the study of variations in genes encoding drug targets that can affect target protein function, a drug-target interaction that produces profound effects on drug response. As noted by Sir William Osler (1849–1919):

> *Variability is the law of life, and as no two faces are the same, so no two bodies are alike, and no two individuals react alike, and behave alike under the abnormal conditions which we know as disease.*

The following sections describe some examples of the importance of pharmacodynamics in the genetic determination of drug response and toxicity.

18.3.1 Pharmacodynamics and Targets for Drug Action

The effect of a drug is a result of their interaction with the receptor, which alters its function and initiates a cascade of physiological and biochemical changes. Here we look at some examples of genetic variations in drug targets that affect their pharmacodynamics.

18.3.1.1 Statins

Statins are competitive inhibitors of hydroxymethylglutaryl-coenzyme A (HMG-CoA) reductase, which is the enzyme that controls the rate-limiting step of cholesterol synthesis. While the majority of patients on statins respond with a 30–50% LDL cholesterol reduction, there is wide (10–70%) variability in response. Genetic variation can result in alternative splicing of *HMGCR*, resulting in reduced LDL lowering, as the variant *HMGCR* is less sensitive to statins [3–5]. (See Chapter 31).

18.3.1.2 Antiplatelet Drugs

Aspirin

Aspirin is well established for primary and secondary prevention of cardiovascular diseases. It causes irreversible acetylation of serine at position 530 of platelet cyclooxygenase-1 (COX-1), the enzyme that catalyzes the conversion of arachidonic acid to prostaglandin G2 (PGG2). This procss is followed by conversion to thromboxane A2 (TXA2), which functions as a potent vasoconstrictor and platelet activator through binding to the TXA2-specific platelet receptor. With adequate dosing and compliance, aspirin is capable of completely inhibiting COX-1 in >99% of patients, and platelets exposed to aspirin are essentially permanently disabled from generating TXA2 and are inhibited throughout their lifespan (approximately 7–10 days) [6]. Even though aspirin disables COX-1-dependent activation, platelet activation can proceed along pathways that are COX-1 independent or only partly dependent. Alternate agonists such as adenosine diphosphate and collagen can produce robust aggregation in the face of complete COX-1 inhibition [7]. This phenotype demonstrates characteristics that make it attractive for pharmacogenetic studies—wide interindividual variability [8], heritability ($h^2 = 0.27 - 0.76$) [9], and association with future cardiovascular events at high levels [10].

The term "aspirin resistance" has been used to describe the lack of pharmacologic drug effects measured by various platelet function tests (e.g., bleeding time, measurement of thromboxane levels and aspirin- or collagen-induced platelet aggregation) or the occurrence of recurrent atherothrombotic events despite drug treatment. Because aspirin uniformly inhibits its target, COX-1, there are probably no pharmacogenetic effects on PK/PD. Studies of genetic variants in the COX-1 gene and indirect COX-1 pathways (*PEAR1*, integrin β3, *GPVI*, *F2R*, and *GP1BA*) have not been validated, and the results are inconclusive [6]. In summary, there is currently no role for genetic testing for aspirin.

Clopidogrel

Thienopyridines are effective antiplatelet agents in patients with acute coronary syndrome (ACS) and percutaneous coronary intervention (PCI). Platelet function in response to clopidogrel is variable [11] and heritable ($h^2 = 0.73$) [12], indicating a genetic basis. Clopidogrel pharmacogenetics has identified drug response variants in genes affecting both its pharmacokinetics (*CYP2C19*, *CYP2C9*, *ABCB1*, *PON1*) and pharmacodynamics (*P2RY12*). A clear relationship has been established between variants in *CYP2C19* and clopidogrel efficacy in terms of platelet inhibition and serious adverse cardiovascular (CV) events [13]. However, one candidate gene study reported that carriers of a common *P2YR12* haplotype (rs6798347, rs6787801, rs9859552, rs6801273, rs9848789, and rs2046934) experienced improved platelet inhibition, although this has not been validated [14]. Unlike the variants in *CYP2C19*, those in *P2RY12* have not been consistently linked to clinical outcomes [15–17]. (See Chapter 25.)

Beta-Blockers

The pharmacological effects of beta-blockers derive from their ability to antagonize beta adrenergic receptors (β-ARs). Thus, the genes for these receptors have been a primary focus in β-blocker pharmacogenetic studies. In the cardiovascular system, there are two β-ARs that beta-blockers can antagonize: β1-AR and β2-AR, both of which are members of the G-protein–coupled receptors superfamily. The β1-AR gene (*ADRB1*) is located on chromosome 10q24–26. It encodes a 51.3 kDa protein, with 477 amino acid residues. β1-ARs are primarily found in the heart, controlling contractility and heart rate, but they are also expressed in kidney, vasculature, and adipose tissues. β2-ARs (*ADRB2*) are more widely distributed, including heart, vasculature, respiratory smooth muscle, kidney, adipose tissue, and brain. The gene consists of one exon located on chromosome 5q31–q32.

Beta-blockers are primarily used in the management of hypertension, heart failure, and coronary artery disease (see Chapter 32). Response to these drugs is highly variable with 30−60% of hypertensive patients on β-blocker monotherapy failing to achieve adequate blood pressure control [18] and approximately 25% of patients discontinuing the drug. Because of intolerance. Ethnic differences in dose response for propranolol [19] motivated pharmacodynamic genetic investigations of beta-blockers.

Variations in *CYP2D6* (pharmacokinetic) and *ADRB1*, *ADRB2*, and *GRK5* (pharmacodynamic) have received the most attention. Two variants in *ADBR1*, Ser49Gly (rs1801252) and Arg389Gly (rs1801253), lead to impaired downregulation [20] and higher signal transduction, respectively [21]. Therefore, carriers of either variant have enhanced β1-receptor activity and more beta-blocker sensitivity. Healthy volunteers and patients with hypertension who carry two Arg389 variants see a greater reduction in HR [22] or BP [23] mainly with metoprolol, although not with all beta-blockers [18] In patients with systolic heart failure treated with either metoprolol or carvedilol [24], but not bucindolol [25], carriers with two copies of the *ADBR1* Arg389 variant had significantly greater improvements in LVEF compared to the Gly389 carriers [26].

In general, carriers of the Arg389 variant have (1) enhanced HR and BP reduction. (2) larger improvements in LVEF, and (3) longer survival when treated with chronic beta-blocker therapy compared to carriers of the Gly389 variant. Although it is unlikely that beta-blocker therapy will ever be withheld for carriers of Gly389, a potential application of these findings is to consider advanced heart failure therapies (e.g., left ventricular assist devices, biventricular pacing, or transplantation) at an earlier stage in patients with the Gly389 variant [26].

Downstream of the beta-1-adrenergic receptor are G-protein–coupled receptor kinases, which are responsible for desensitization of the β1-adrenergic receptor. A Glu41Leu genetic variant in G protein–coupled receptor kinase 5 (*GRK5*), is more prevalent in African Americans. GLULeu41 resulting in more effective uncoupling of isoproterenol-stimulated responses than *GRK5*-Q41, thus producing a pharmacological-like “beta-blockade” in mice [27], albeit with no differences in atenolol-induced HR reduction in humans [28].

Because certain beta-blockers such as atenolol and carvedilol are minimally handled by CYP2D6, they may be reasonable alternates for *CYP2D6**4 carriers with metoprolol-induced bradycardia [26].

ADRB2 and Asthma

Asthma is a complex, chronic inflammatory airways disease characterized by airway hypersensitivity and airflow obstruction. However, clinical presentation can be acute and sometimes life threatening. Asthma has a known genetic link and is a widely studied disease in terms of genetics, but it has a complex etiology with interaction between genetics and environmental factors, including allergens and infections [29]. (See Chapter 25) [4].

β2 adrenoceptors are essential drug targets in asthma therapy, with inhaled long- and short-acting β2 adrenoceptor agonists being among the common prescribed therapies. The β2 adrenoceptor mediates the responsiveness of the airways, including bronchodilation, mucociliary clearance, and cholinergic neurotransmission inhibition [30]. Patients show a wide variation in response to both inhaled long-acting β2-adrenoceptor agonists (LABA) and short-acting

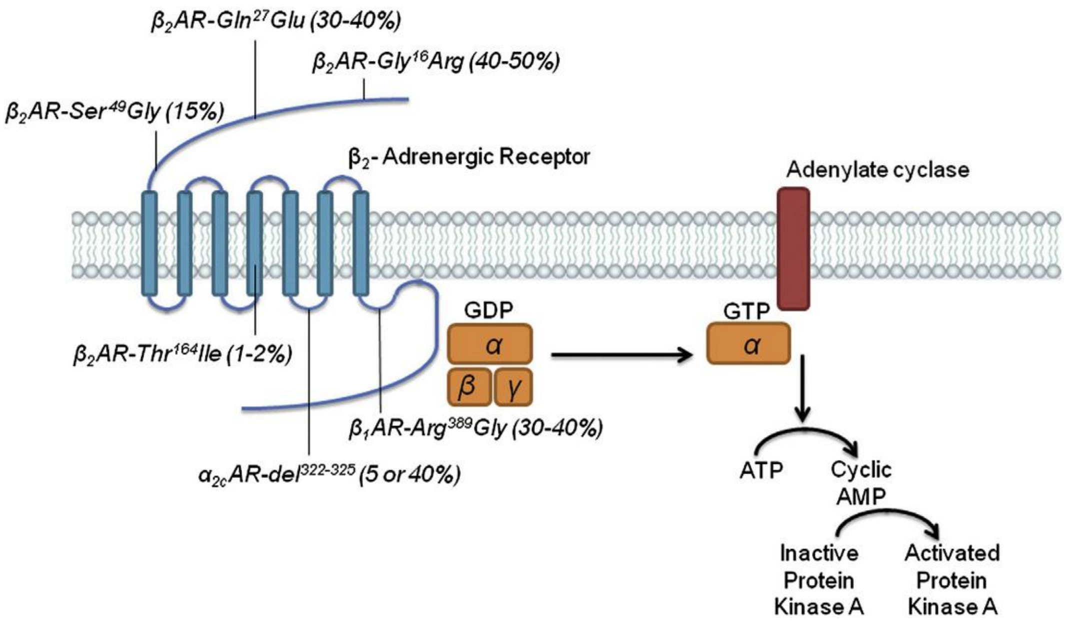

β_2AR-Gln^{27}Glu (30-40%)
β_2AR-Gly^{16}Arg (40-50%)
β_2AR-Ser^{49}Gly (15%)
β_2- Adrenergic Receptor
Adenylate cyclase
GDP
GTP
α
β
γ
α
β_2AR-Thr^{164}Ile (1-2%)
β_1AR-Arg^{389}Gly (30-40%)
α_{2c}AR-del$^{322\text{-}325}$ (5 or 40%)
ATP
Cyclic AMP
Inactive Protein Kinase A
Activated Protein Kinase A

FIGURE 18.9 Structure of the β2 adrenoceptor.

β2 adrenoceptor agonists (SABA). Structure of the β2 adrenoceptor is shown in Figure 18.9. The *ADRB2* gene is located on the long arm of chromosome 5, a region that positively associates with asthma [31]. More than 50 genetic polymorphisms in the *ADRB2* gene have been identified that alter receptor behavior after agonist exposure, and variation may determine the extent of bronchodilator response.

A coding single-nucleotide polymorphism (SNP) in the *ADRB2* gene that results in the substitution of glycine for arginine at the 16th amino acid (Gly-16) has been associated with a more severe asthma phenotype and with variation in response to inhaled β2 adrenoceptor agonists [31,32].

In one study, children who were homozygous for Arg16 were 5.3 times more likely to show acute reversibility to SABA than those homozygous for Gly16 with heterozygous Gly16/Arg16, who show an intermediate response [33]. In another study, subjects with asthma who were homozygous for Arg16 and treated with regular SABA experienced a significant decline in morning and night time peak expiratory flow (PEF) compared with subjects who were homozygous for Gly16 [31,33,34]. There have been conflicting studies related to the Gly16 SNP and LABA use [31,35,36] (see Chapter 28).

HIV and *CCR5*

HIV infection in humans involves C-C chemokine receptor type 5 (*CCR5*, *CD195*), which is used by HIV as a coreceptor for entry into target cells, and *CCR2* (*CD192*), which is for monocyte chemoattractant protein 1. A polymorphism of *CCR2* at codon 64 (V64I) is common with an allele frequency of 10%, and 9% of Caucasians carry a 32-base pair deletion (*CCR5* Δ32) in the *CCR5* gene. The *CCR2* V64I allele has a significant impact on AIDS progression and *CCR5* Δ32 protects against HIV-1 transmission [37]. A HIV integrase inhibitor, such as maraviroc, specifically interferes with the interaction between *CCR5* and HIV. However, studies have shown no difference in virologic response to maraviroc-based antiretroviral according to *CCR5* Δ32 genotype in 982 patients genotyped in MOTIVATE (Maraviroc versus Optimized Therapy in Viremic Antiretroviral Treatment-experienced Patients) trials [38].

Cancer

Examples from cancer chemotherapy highlight the importance of pharmacodynamic genetic variation in drug response, and focused profiling is now becoming a part of routine cancer care [39] (see Chapters 20, 21, and 34). KRAS is a membrane GTPase that can activate many proteins in EGFR signaling pathways such as c-Raf and PI3K. Aberrant activation of these proteins usually causes cancer development independent of upstream EGFR signaling. Cetuximab and panitumumab are two monoclonal antibodies that were designed to inhibit the growth and survival of tumor cells with overexpressed EGFR in colon and head and neck cancers. If KRAS is actively mutated, the inactivation of EGFR by cetuximab or panitumumab will have no beneficial effect in curing KRAS-induced cancers [40]. The FDA label for these drugs recommends use only in patients with EGFR-expressing colon cancer who are KRAS-mutant negative (wild-type) [41].

Crizotinib is an *ALK* inhibitor approved to treat non-small-cell lung cancer (NSCLC). *ALK* was found to form a fusion *EML4–ALK* gene in 3–5% of NSCLC patients. This fusion gene is a constitutively activated kinase and leads to carcinogenesis. Therefore, its detection is required for the crizotinib use. In clinical treatment of NSCLC caused by *EML4–ALK*, drug resistance is linked to several tumor-specific genetic mutations in *ALK*: L1196M, C1156Y, F1174L, and G1269A [42,43]. Moreover, the FDA has recommends *ALK* positive as the indicator for prescribing crizotinib to treat patients with locally advanced or metastatic NSCLC [44]. In acute lymphoblastic leukaemia, the presence of a 9:22 translocation is an indication for imatinib. Breast cancer with amplification of HER2 shows improved response with the anti-HER2 monoclonal antibody trastuzumab or the HER2 tyrosine kinase inhibitor lapatinib. Lung cancer, melanoma, and myeloproliferative disorders tend to be sensitive to tyrosine kinase inhibitors with mutations in the respective genes *EGFR*, *BRAF*, and *JAK2* [45].

Cystic Fibrosis CFTR Mutations and Ivacaftor

Cystic fibrosis is a multisystem autosomal recessive-genetic condition affecting 1 in 2500 births in children of European decent and 1 in 32000 children of Asian decent. For people of European descent, 1 in 25 carry a cystic fibrosis mutation while 1 in 90 people of Asian descent are carriers. Cystic fibrosis results in infantile failure to thrive, accumulation of viscous mucus within the respiratory tract and recurrent infection, loss of exocrine pancreatic function, impaired intestinal absorption, infertility, and an increase in sweat chloride concentration. Treatment is supportive care, prevention and treatment of respiratory infection, management of diabetes and dietary advice, and ultimately lung and or pancreas transplant.

Cystic fibrosis is characterized by the loss of epithelial chloride transport due to mutations in the cystic fibrosis transmembrane-conductance regulator gene that encodes for the CFTR protein. The CFTR protein is a chloride channel that is usually present at the cell surface of epithelial cells [10]. The CFTR chloride channel is opened and closed by ATP binding and hydrolysis when activated by protein kinase A. CFTR normally transports chloride to regulate salt, fluid, and acid-base balance throughout the body [46]. There are more than 1900 known genetic CFTR mutations, each with varying degrees of phenotypic severity [46]. F508del is by far the most common mutation, accounting for two-thirds of cases.

CFTR is composed of two transmembrane-spanning domains that form an ion channel, two nucleotide-binding domains that form the ATP binding targets, and a regulatory domain that contains the PKA phosphorylation site [47]. CFTR structure and associated mutations are shown in Figure 18.10.

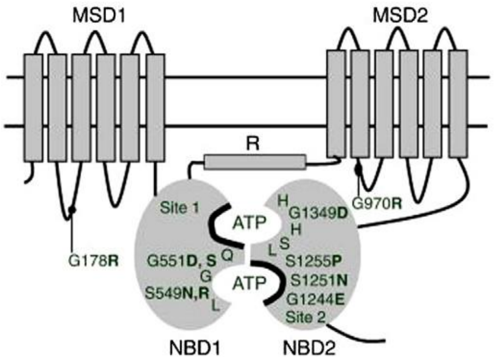


FIGURE 18.10 Structure of CFTR and associated mutations [47].

Ivacaftor is a CFTR potentiator that increases CFTR channel activity and has been shown to improve lung function in children 6 years and older with cystic fibrosis who have the G551D-CFTR gating mutation. Randomized controlled trials have shown an increase in FEV1, a decrease in pulmonary exacerbations, and a decrease in sweat chloride in patients with cystic fibrosis with at least one copy of the G551D mutation [48–50]. There was no effect seen in patients with the common F508del mutation.

Thymidylate Synthase

Thymidylate synthase (TS) is an enzyme that catalyzes the conversion of deoxy-uridylate to deoxy-thymidylate and is essential for the synthesis of DNA [51]. The human TS gene is polymorphic with either double or triple tandem repeats of the 28 base-pair sequences downstream of the cap-site in the 5′-terminal regulatory region [52]. 5-fluorouracil (5-FU) is an anti-cancer drug used in the treatment of colonic and gastric cancer that acts through its active metabolite 5-flurodeoxyuridylate, leading to the inhibition of TS. Individuals categorized as homozygous for the double repeat variant (SS) had a 50% response rate to 5-FU therapy compared to 9% in individuals characterized as homozygous for the triple repeat variant (LL) and 15% in those with the SL genotype [53]. In addition, patients with the LL genotype had less severe side effects. This indicates that genotyping patients for the TS polymorphism may lead to the ability to identify those who are more likely to benefit from 5-FU-based chemotherapy and those who are more likely to suffer from adverse drug reactions.

18.3.2 Genetic Variability that Indirectly Alters Drug Response

Changes in drug response can be a result of genetic variability, which is not directly a result of changes in drug targets or transporters and may result in altered drug efficacy or increased susceptibility to adverse drug reactions.

18.3.2.1 Glucose-6-Phosphate Dehydrogenase Deficiency

Glucose-6-phosphate dehydrogenase (G6PD) deficiency was first described in 1956 following the development of hemolytic anemia in patients receiving primaquine therapy [54]. G6PD deficiency is the most common enzyme defect in humans, affecting approximately 400 million worldwide [55]. There are marked ethnic and geographical differences in G6PD prevalence, from <1 in 1000 in Northern Europeans to 1 in 2 Kurdish Jews. G6PD deficiency is an X-linked disorder. Although it was initially assumed to be a single-gene disorder, there are now more than 300 different known mutations. The frequency of G6PD deficiency in certain geographical areas is high because affected individuals and heterozygous female carriers have a selective survival advantage against severe malaria [55].

G6PD plays an essential role in preserving the integrity of the erythrocyte and is central in the defense against oxidative stress. The human G6PD molecule (Figure 18.11) is a dimer that provides NADPH for the antioxidant cascade in the erythrocyte [55]. The active form catalyzes the first step in the pentose phosphate pathway (PPP): the oxidation of glucose-6-phosphate to 6-phosphogluconolactone with a concomitant reduction of NAPD to NADPH [55]. NADPH is integral in protecting the erythrocyte against oxidative damage by maintaining a high level of reduced glutathione (GSH) in the cell [55]. GSH prevents oxidation of the sulfhydryl groups in hemoglobin and in the erythrocyte membrane [55]. G6PD is the only NADPH-producing enzyme that can be activated in response to oxidative stress and therefore to G6PD deficiency. NADPH concentrations are low, leading to low levels of GSH that result in oxidative damage, leading to hemolysis [55]. The pathway can be seen in Figure 18.12.

Drug-induced hemolytic anemia is a common consequence of G6PD deficiency. Many drugs have been identified as causing or potentially causing hemolysis in patients with it. Some drugs, including seven that are commonly used, are contraindicated in G6PD deficiency: dapsone, methylene blue, nitrofurantoin, primaquine, rasburicase, tolonium chloride, and phenazopyridine [56].

18.3.2.2 Factor V, Prothrombin, and the Oral Contraceptive Pill

It is well recognized that some women who take the oral contraceptive pill (OCP) are at increased risk of developing deep vein thrombosis (DVT) and pulmonary thrombosis/thromboembolism (PTE). Genetic variations associated with prothrombin and factor V have been identified and are thought to be responsible for increased risk of thrombosis in susceptible women taking OCPs [57].

Thrombin converts factor V to its active form, factor Va, which is essential for the coagulation cascade and acts as a co-factor for the conversion of prothrombin to thrombin by factor Xa. Activated protein C (APC) is a potent anticoagulant that limits coagulation by proteolytic factor Va inactivation [58]. The coagulation cascade is summarized in Chapter 23. APC resistance is associated with a seven-fold increased risk of DVT [59]. The phenotype of APC resistance is associated with the heterogenicity and homogenicity of a SNP found in exon 10 of the factor V gene (nucleotide position 1691, G to A substitution), leading to a factor V molecule (FV Leiden or FV Q506) that is not fully inactivated by APC [59].

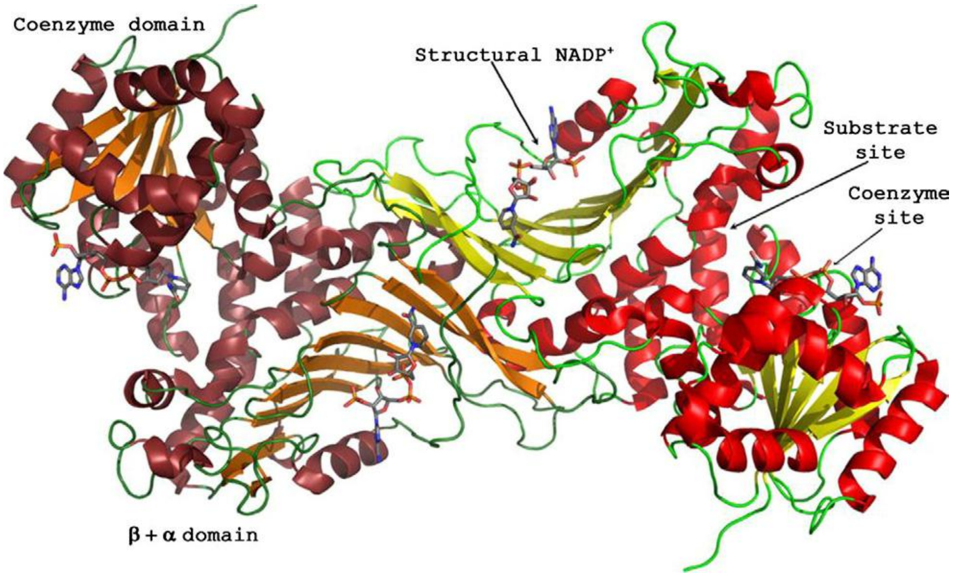


FIGURE 18.11 Human G6PD molecule [55].

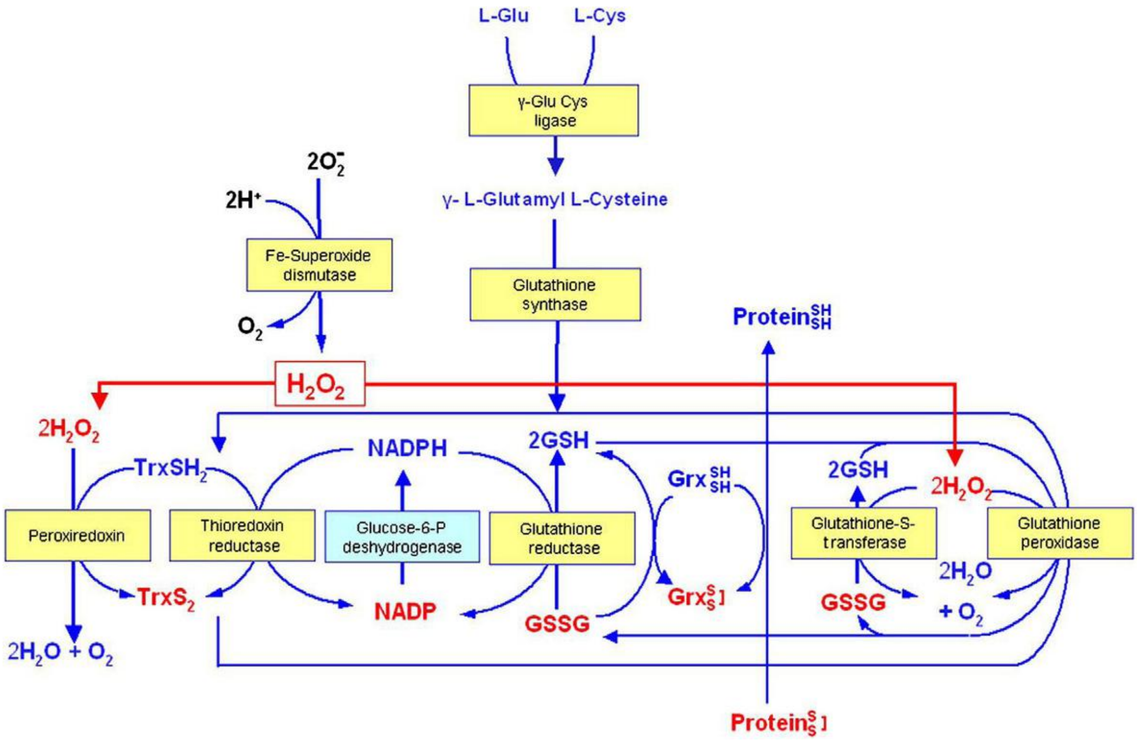

L-Glu
L-Cys
γ-Glu Cys ligase
γ- L-Glutamyl L-Cysteine
$2O_2^-$
$2H^+$
Fe-Superoxide dismutase
O_2
Glutathione synthase
$Protein^{SH}_{SH}$
H_2O_2
$2H_2O_2$
$TrxSH_2$
NADPH
2GSH
Grx^{SH}_{SH}
2GSH
$2H_2O_2$
Peroxiredoxin
Thioredoxin reductase
Glucose-6-P deshydrogenase
Glutathione reductase
Glutathione-S-transferase
Glutathione peroxidase
$TrxS_2$
NADP
GSSG
$Grx^{S}_{S}]$
GSSG
$2H_2O$
$+ O_2$
$2H_2O + O_2$
$Protein^{S}_{S}]$

FIGURE 18.12 G6PD pathway [55].

The estrogen-containing OCP induces changes in procoagulant proteins and natural anticoagulants, leading to a state of hypercoagulability with increased fibrin production. Plasma levels of prothrombin and fibrinogen increase while factor V decreases; however, the most important effect may be acquired APC resistance [58]. These changes are evident within three months of commencing the OCP and disappear four months after withdrawal [58]. The use of the OCP by women with factor V Leiden is associated with a 17× higher risk of venous thromboembolism compared with women without factor V Leiden using no contraception [60].

A SNP in the gene encoding for prothrombin (A20210, G to A substitution) was identified as being associated with an three-fold increased risk of venous thromboembolism [61]. The risk of DVT in patients with A20210 using the OCP was increased six-fold [61].

18.3.3 Adverse Drug Reactions

An adverse drug reaction (ADR) is an unwanted or harmful reaction following administration of a medication or a combination of medications. It is suspected to be related to the drug administered. ADRs are a significant cause of morbidity and mortality worldwide [62]. They are also a significant financial burden for healthcare systems and lead to longer hospital stays [62]. Approximately 40% of patients in a community setting experience ADRs, while 6.5% of hospital admissions of adults and 2.1% of those of children are related to ADRs [62,63]. Of hospitalized patients, 6.7% suffer serious ADRs [64].

There are many examples of ADRs related to known genetic polymorphisms, and the study of these associations is often termed toxigenomics or toxigenetics. Understanding them will promote stratified medicine in the future. Generally the genes that influence drug toxicity relate to drug-metabolizing enzymes, drug transporters, and human leukocyte antigens (HLAs); see Chapters 20 and 21. Malignant hyperthermia is one example of an ADR that is related to a genetic variation of a drug target. It is discussed next.

18.3.3.1 Malignant Hyperthermia

Malignant hyperthermia is a serious and potentially fatal inherited disorder of skeletal muscle metabolism that was identified in 1985. It is triggered by commonly used volatile inhalation anesthetics and/or depolarizing muscle relaxants.

The incidence and prevalence or malignant hypertension varies greatly between epidemiological studies: from 1 in 50,000 to 1 in 100,000 anaesthetics [65]. Malignant hyperthermia is characterized by abnormalities in excitation-contraction coupling (ECC) and by the dysregulation of myoplasmic calcium levels during the activation of skeletal muscle cells. Clinical features include muscle rigidity, accelerated metabolism, lactic and respiratory acidosis, skeletal muscle breakdown leading to pyrexia, tachycardia, hyperventilation, raised PaCO$_2$, myoglobulinuria, rhabdomyolysis, hyperkalemia, and renal failure.

The ryanodine receptor gene (RYR1) encodes for the key channel that mediates calcium release in skeletal muscle during ECC [65]. Mutations in it, mapping to chromosome 19q13.2, are considered to account for susceptibility to malignant hyperthermia in more than half of cases [66,67]. To date, 80 distinct point mutations in the RYR1 gene have been found to be associated with malignant hyperthermia susceptibility [68].

RYR1 encodes for transcript is >15 kb long; however, most of the mutations linked to malignant hyperthermia are located in residues M1-R614 (MH domain 1), R2162 and R2458 (MH domain 2); a small proportion of mutations associated with MH have been linked to the COOH- terminal hydrophobic region (MH domain 3) [68]. Figure 18.13 shows the mutations identified and the associated domain. RYR1 gene mutations produce abnormalities that alter the channel kinetics, which results in channels that are hyper- and hyposensitive to activating and inactivating ligands. Most malignant hyperthermia–linked mutations cause the channels to be oversensitive [66,68].

Figure 18.14 shows how mutations linked to malignant hyperthermia affect the RYR and thus lead to alterations in calcium homeostasis. RYR1 mutations are thought to be present in 50–70% of families with a history of MH; many of the known mutations have been found only in single families, while others are more common. The G341R mutation is present in approximately 6% of Irish, English, and French families but is rare in Northern Europe, whereas the V2168M mutation is common in Swiss families but is rare elsewhere in the world [67]. RYR1 mutations resulting in increased susceptibility to MH are inherited in an autosomal dominant pattern.

Treatment of MH is with the dantrolene, a smooth muscle relaxant that acts as a ryanodine receptor antagonist. This drug has significantly reduced MH mortality from 80% in the 1970 to <5% in current clinical practice [67]. Patients who may be susceptible to malignant hyperthermia and those with a family history of it can be tested using muscle biopsy and *in vitro* contraction tests or newer tests with DNA analysis. The latter method is limited to approximately 40% of families who are eligible for DNA testing, as only certain mutations may be detected.

18.4 CONCLUSION: PHARMACODYNAMIC PHARMACOGENOMICS INTO CLINICAL PRACTICE

This chapter provides an overview of the principles of pharmacodynamics and the critical role it has in the pharmacogenetic aspects of drug efficacy and adverse reactions. This is reflected in the later chapters of this book,

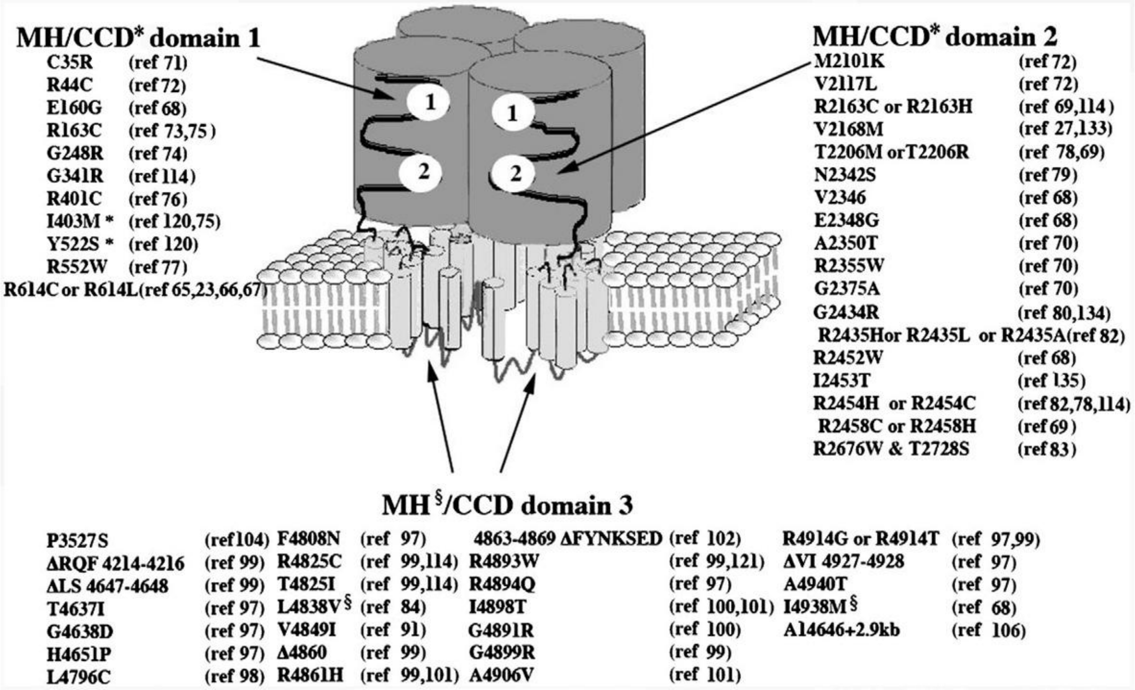

MH/CCD* domain 1
C35R (ref 71)
R44C (ref 72)
E160G (ref 68)
R163C (ref 73,75)
G248R (ref 74)
G341R (ref 114)
R401C (ref 76)
I403M * (ref 120,75)
Y522S * (ref 120)
R552W (ref 77)
R614C or R614L(ref 65,23,66,67)
1
1
2
2
MH/CCD* domain 2
M2101K (ref 72)
V2117L (ref 72)
R2163C or R2163H (ref 69,114)
V2168M (ref 27,133)
T2206M or T2206R (ref 78,69)
N2342S (ref 79)
V2346 (ref 68)
E2348G (ref 68)
A2350T (ref 70)
R2355W (ref 70)
G2375A (ref 70)
G2434R (ref 80,134)
R2435H or R2435L or R2435A(ref 82)
R2452W (ref 68)
I2453T (ref 135)
R2454H or R2454C (ref 82,78,114)
R2458C or R2458H (ref 69)
R2676W & T2728S (ref 83)
MH§/CCD domain 3
P3527S (ref104)
ΔRQF 4214-4216 (ref 99)
ΔLS 4647-4648 (ref 99)
T4637I (ref 97)
G4638D (ref 97)
H4651P (ref 97)
L4796C (ref 98)
F4808N (ref 97)
R4825C (ref 99,114)
T4825I (ref 99,114)
L4838V§ (ref 84)
V4849I (ref 91)
Δ4860 (ref 99)
R4861H (ref 99,101)
4863-4869 ΔFYNKSED (ref 102)
R4893W (ref 99,121)
R4894Q (ref 97)
I4898T (ref 100,101)
G4891R (ref 100)
G4899R (ref 99)
A4906V (ref 101)
R4914G or R4914T (ref 97,99)
ΔVI 4927-4928 (ref 97)
A4940T (ref 97)
I4938M§ (ref 68)
A14646+2.9kb (ref 106)

FIGURE 18.13 **Mutations linked to malignant hyperthermia.** The mutations affect the RyR leading to alterations in calcium homeostasis [68].

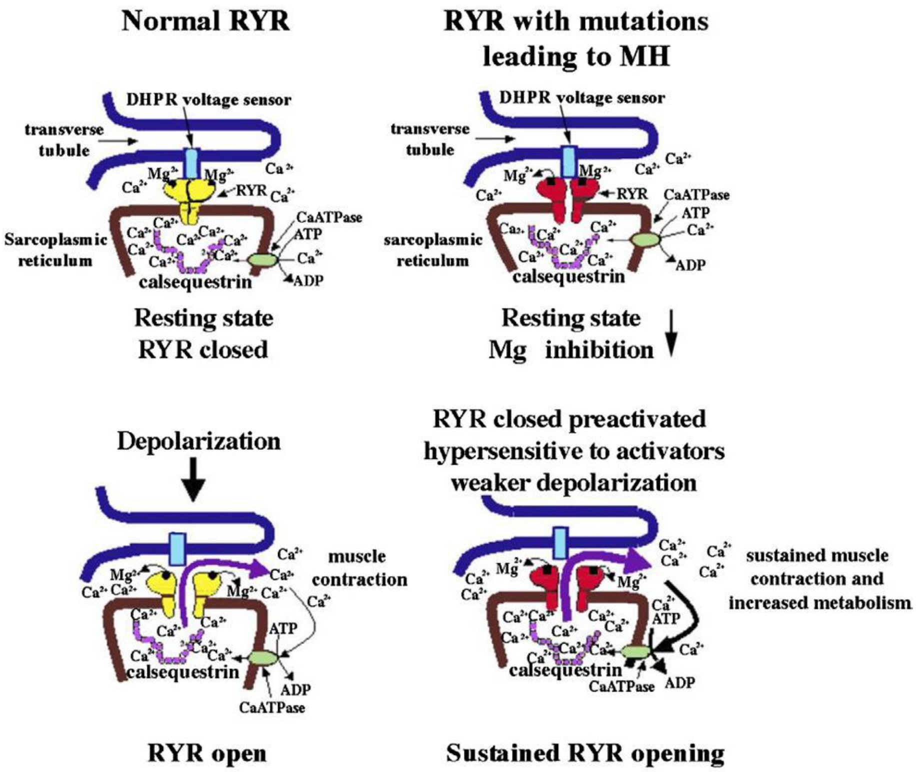

Normal RYR
RYR with mutations leading to MH
DHPR voltage sensor
transverse tubule
Sarcoplasmic reticulum
RYR
CaATPase
ATP
ADP
calsequestrin
Resting state RYR closed
Resting state
Mg inhibition
Depolarization
muscle contraction
RYR closed preactivated hypersensitive to activators weaker depolarization
sustained muscle contraction and increased metabolism
RYR open
Sustained RYR opening

FIGURE 18.14 **Normal RyR receptor and receptor with mutation.** Functional differences of normal and mutant RyR receptor leading to MH [68].

where a majority of pharmacogenetic applications are based on the pharmacodynamic aspects of drugs. A good understanding of pharmacodynamic principles is essential to overcome the challenges in translating the basic science into a cost-effective, point-of-care tool for the everyday clinical setting.

Key Concepts in Pharmacodynamics Pharmacogenomics

1. Pharmacodynamics is the study of the physiological and molecular effects of drugs on their recipient and the relationship between plasma drug concentration and drug effect.
2. Large individual variability in drug efficacy and safety has been known to exist since the beginning of modern medicine; however the reasons for it have proven difficult to elucidate
3. An individual's response to a given drug depends on a complex interplay between environmental and genetic factors.
4. Pharmacodynamic pharmacogenomics is the study of altered responses to drugs that are associated with genetic polymorphisms of genes encoding for drug targets and with modifiers or key proteins.

GLOSSARY

Agonist stimulates or activates the receptor to produce an effect

Allosteric modulation indirect modulation of the effects of an agonist by binding to a separate binding site, leading to an increase or decrease in the receptor's affinity for agonist

Antagonist prevents the action of an agonist but does not have any effect itself

Desensitization repeated exposure to a drug leading to a diminishing effect

Inverse agonist produces an opposite effect to that of the agonist

Organ bath standard for *in vitro* pharmacodynamic dose-response experiments involving suspension of an isolated section of organ/tissue in a controlled solution and sensors to measure the contraction or relaxation of the tissue when drugs are added

Partial agonist stimulates the receptor to a limited extent but also prevents any further stimulation by naturally occurring agonists

Pharmacodynamics the study of the physiological and molecular effects of drugs on their recipient and the relationship between plasma drug concentration and drug effect

Polymorphism a variation in DNA sequence that is present at an allele frequency of $\geq 1\%$ of the population

Receptor a component of the organism with which a drug interacts

Second messenger a molecule that relays signals from receptors to the cell surface target

SNP single-nucleotide polymorphism—a point mutation occurring at a frequency of $\geq 1\%$

Tachyphylaxis decrease in response to a drug following repeated administration

Therapeutic index the relationship between therapeutic and toxic doses of a drug

ACRONYMS AND ABBREVIATIONS

5-FU	5-fluorouracil
ACE	angiotensin-converting enzyme
ADR	adverse drug reaction
APC	activated protein C
Bmax	maximum specific binding
cAMP	cyclic adenosine monophosphate
COX	cyclo-oxygenase
CFTR	cystic fibrosis transmembrane conductance regulator
cGMP	cyclic guanosine monophosphate
DAG	diacylglycerol
DNA	deoxyribonucleic acid
DVT	deep venous thrombosis
EC	extracellular
ECC	excitation-contraction coupling
ED$_{50}$	therapeutically effective dose in 50% of people given drug
G6PD	glucose 6 dehydrogenase deficiency
GABA	gamma-aminobutyric acid
GPCR	G-protein coupled receptor
HLA	human leukocyte antigen
IC	intracellular
IUPHAR	International Union of Basic and Clinical Pharmacology
KD$_{50}$	agonist dissociation constant
LABA	long-acting β2 adrenoceptor agonist
NADPH	nicotinamide adenine dinucleotide phosphate
OCP	oral contraceptive pill
PPAR	peroxisome proliferator-activated receptor
PET	positron emission tomography
PPP	pentose phosphate pathway
PTE	pulmonary thrombosis/thromboembolism
RYR1	ryanodine receptor gene
SABA	short-acting β2 adrenoceptor agonist
SNP	single-nucleotide polymorphism
TI	therapeutic Index
TS	thymidylate synthase
VKORC1	vitamin K epoxide reductase complex
WHO	World Health Organization

REFERENCES

[1] Brunton LL, Blumenthal DK, Murri N, Dandan RH, Knollmann BC. Goodman & Gilman's the Pharmacological Basis of Therapeutics. 12th ed. McGraw-Hill; 2011.

[2] Sharman JL, Benson HE, Pawson AJ, Lukito V, Mpamhanga CP, Bombail V, NC-IUPHAR. IUPHAR-DB: updated database content and new features. Nucl Acids Res 2013;41(Database issue):D1083–808.

[3] Chasman DI, Posada D, Subrahmanyan L, Cook NR, Stanton Jr VP, Ridker PM. Pharmacogenetic study of statin therapy and cholesterol reduction. JAMA 2004;291(23):2821–7.

[4] Krauss RM, Mangravite LM, Smith JD, Medina MW, Wang D, Guo X, et al. Variation in the 3-hydroxyl-3-methylglutaryl coenzyme a reductase gene is associated with racial differences in low-density lipoprotein cholesterol response to simvastatin treatment. Circulation 2008;117(12):1537–44.

[5] Medina MW, Gao F, Ruan W, Rotter JI, Krauss RM. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. Circulation 2008;118(4):355–62.

[6] Faraday N, Becker DM, Becker LC. Pharmacogenomics of platelet responsiveness to aspirin. Pharmacogenomics 2007;8(10):1413–25.

[7] Gurbel PA, Bliden KP, DiChiara J, Newcomer J, Weng W, Neerchal NK, et al. Evaluation of dose-related effects of aspirin on platelet function: results from the aspirin-induced platelet effect (ASPECT) study. Circulation 2007;115(25):3156–64.

[8] Zufferey A, Reny JL, Combescure C, de Moerloose P, Sanchez JC, Fontana P. Platelet reactivity is a stable and global phenomenon in aspirin-treated cardiovascular patients. Thromb Haemost 2011;106(3):466–74.

[9] Faraday N, Yanek LR, Mathias R, Herrera-Galeano JE, Vaidya D, Moy TF, et al. Heritability of platelet responsiveness to aspirin in activation pathways directly and indirectly related to cyclooxygenase-1. Circulation 2007;115(19):2490–6.

[10] Frelinger 3rd AL, Li Y, Linden MD, Barnard MR, Fox ML, Christie DJ, et al. Association of cyclooxygenase-1-dependent and -independent platelet function assays with adverse clinical outcomes in aspirin-treated patients presenting for cardiac catheterization. Circulation 2009;120(25):2586–96.

[11] Gurbel PA, Bliden KP, Guyer K, Cho PW, Zaman KA, Kreutz RP, et al. Platelet reactivity in patients and recurrent events post-stenting: results of the PREPARE POST-STENTING Study. J Am Coll Cardiol 2005;46(10):1820–6.

[12] Shuldiner AR, O'Connell JR, Bliden KP, Gandhi A, Ryan K, Horenstein RB, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. JAMA 2009;302(8):849–57.

[13] Scott SA, Sangkuhl K, Stein CM, Hulot JS, Mega JL, Roden DM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C19 Genotype and clopidogrel therapy: 2013 update. Clin Pharmacol Ther 2013;94(3):317–23.

[14] Rudez G, Bouman HJ, van Werkum JW, Leebeek FW, Kruit A, Ruven HJ, et al. Common variation in the platelet receptor P2RY12 gene is associated with residual on-clopidogrel platelet reactivity in patients undergoing elective percutaneous coronary interventions. Circ Cardiovasc Genet 2009;2(5):515–21.

[15] Simon T, Verstuyft C, Mary-Krause M, Quteineh L, Drouet E, Méneveau N. French registry of acute ST-elevation and non-ST-elevation myocardial infarction (FAST-MI) investigators. Genetic determinants of response to clopidogrel and cardiovascular events. N Engl J Med 2009;360(4):363–75.

[16] Angiolillo DJ, Fernandez-Ortiz A, Bernardo E, Ramírez C, Cavallari U, Trabetti E, et al. Lack of association between the P2Y12 receptor gene polymorphism and platelet response to clopidogrel in patients with coronary artery disease. Thromb Res 2005;116(6):491–7.

[17] Cuisset T, Frere C, Quilici J, Morange PE, Saut N, Lambert M, et al. Role of the T744C polymorphism of the P2Y12 gene on platelet response to a 600-mg loading dose of clopidogrel in 597 patients with non-ST-segment elevation acute coronary syndrome. Thromb Res 2007;120(6):893–9.

[18] Shin J, Johnson JA. Pharmacogenetics of beta-blockers. Pharmacotherapy 2007;27(6):874–87.

[19] Venter CP, Joubert PH. Ethnic differences in response to beta 1-adrenoceptor blockade by propranolol. J Cardiovasc Pharmacol 1984;6(2):361–4.

[20] Rathz DA, Brown KM, Kramer LA, Liggett SB. Amino acid 49 polymorphisms of the human beta1-adrenergic receptor affect agonist-promoted trafficking. J Cardiovasc Pharmacol 2002;39(2):155–60.

[21] Mason DA, Moore JD, Green SA, Liggett SB. A gain-of-function polymorphism in a G-protein coupling domain of the human beta1-adrenergic receptor. J Biol Chem 1999;274(18):12670–4.

[22] Liu J, Liu ZQ, Tan ZR, Chen XP, Wang LS, Zhou G, et al. Gly389Arg polymorphism of beta1-adrenergic receptor is associated with the cardiovascular response to metoprolol. Clin Pharmacol Ther 2003;74(4):372–9.

[23] Johnson JA, Zineh I, Puckett BJ, McGorray SP, Yarandi HN, Pauly DF. Beta 1-adrenergic receptor polymorphisms and antihypertensive response to metoprolol. Clin Pharmacol Ther 2003;74(1):44–52.

[24] Mialet Perez J, Rathz DA, Petrashevskaya NN, Hahn HS, Wagoner LE, Schwartz A, et al. Beta 1-adrenergic receptor polymorphisms confer differential function and predisposition to heart failure. Nat Med 2003;9(10):1300–5.

[25] Liggett SB, Mialet-Perez J, Thaneemit-Chen S, Weber SA, Greene SM, Hodne D, et al. A polymorphism within a conserved beta(1)-adrenergic receptor motif alters cardiac function and beta-blocker response in human heart failure. Proc Natl Acad Sci USA 2006;103(30):11288–93.

[26] Shin J, Johnson JA. Beta-blocker pharmacogenetics in heart failure. Heart Fail Rev 2010;15(3):187–96.

[27] Liggett SB, Cresci S, Kelly RJ, Syed FM, Matkovich SJ, Hahn HS, et al. A GRK5 polymorphism that inhibits beta-adrenergic receptor signaling is protective in heart failure. Nat Med 2008;14(5):510–7.

[28] Kurnik D, Cunningham AJ, Sofowora GG, Kohli U, Li C, Friedman EA, et al. GRK5 Gln41Leu polymorphism is not associated with sensitivity to beta(1)-adrenergic blockade in humans. Pharmacogenomics 2009;10(10):1581–7.

[29] <www.who.int/topics/asthma/en/>. [accessed 10.07.2013], 2040.

[30] Hizawa N. Beta-2 adrenergic receptor genetic polymorphisms and asthma. J Clin Pharm Ther 2009;34(6):631–43.

[31] Postma DS, Bleecker ER, Amelung PJ, Holroyd KJ, Panhuysen CI, Meyers DA, et al. Genetic susceptibility to asthma-bronchial hyperresponsiveness coinherited with a major gene for atopy. N Engl J Med 1995;333:894–900.

[32] Contopoulos-Ioannidis DG, Manoli EN, Ioannidis JP. Meta analysis of the associate of beta-2-adrenergic receptor polymorphism with asthma phenotypes. J Allergy Clin Immunol 2005;115:963–72.

[33] Martinez FD, Graves PE, Baldini M, Solomon S, Erickson R. Association between genetic polymorphisms of the beta2-adrenoceptor and response to albuterol in children with and without a history of wheezing. J Clin Invest 1997;100:3184–8.

[34] Israel E, Drazen JM, Liggett SB, Boushey HA, Cherniack RM, Chinchilli VM, et al. The effect of polymorphisms of the beta(2) adrenergic receptor on the response to regular use of albuterol in asthma. Am J Resp Crit Care 2000;162(1):75–80.

[35] Israel E, Chinchilli VM, Ford JG, Boushey HA, Cherniack R, Craig TJ, et al. Use of regularly scheduled albuterol treatment in asthma: genotype stratified, randomized placebo-controlled crossover trial. Lancet 2004;364:1505–12.

[36] Wechsler ME, Lehman E, Lazarus SC, Lemanske RF, Boushey HA, Deykin A, et al. Beta-adrenergic receptor polymorphisms and response to salmeterol. Am J Resp Crit Care 2006;173:519–26.

[37] Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature 1996;382(6593):722–5.

[38] Fätkenheuer G, Nelson M, Lazzarin A, Konourina I, Hoepelman AI, Lampiris H, et al. MOTIVATE 1 and MOTIVATE 2 Study Teams. Subgroup analyses of maraviroc in previously treated R5 HIV-1 infection. N Engl J Med 2008;359(14):1442–55.

[39] McLeod HL. Cancer pharmacogenomics: early promise, but concerted effort needed. Science 2013;339(6127):1563–6.

[40] Lièvre A, Bachet JB, Le Corre D, Boige V, Landi B, Emile JF, et al. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res 2006;66(8):3992–5.

[41] Kranenburg O. The KRAS oncogene: past, present, and future. Biochim Biophys Acta 2005;1756(2):81–2.

[42] Doebele RC, Pilling AB, Aisner DL, Kutateladze TG, Le AT, Weickhardt AJ, et al. Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. Clin Cancer Res 2012;18(5):1472–82.

[43] Zhang S, Wang F, Keats J, Zhu X, Ning Y, Wardwell SD, et al. Crizotinib-resistant mutants of EML4-ALK identified through an accelerated mutagenesis screen. Chem Biol Drug Des 2011;78(6):999–1005.

[44] Pillai RN, Ramalingam SS. The biology and clinical features of non-small cell lung cancers with EML4-ALK translocation. Curr Oncol Rep 2012;14(2):105–10.

[45] Welch BM, Kawamoto K. Clinical decision support for genetically guided personalized medicine: a systematic review. J Am Med Inform Assoc 2013;20(2):388–400.

[46] Van Goor F, Yu H, Burton B, Hoffman BJ. Effect of ivacaftor on CFTR forms with missense mutation associated with defects in protein processing or function. J Cyst Fibros 2013;7:1–8.

[47] Yu H, Huang CJ, Worley J, Cao D, Johnson JP, Urrutia A, et al. Ivacaftor potentiation of multiple CFTR channels with gating mutations. J Cyst Fibros 2012;11(3):237–45.

[48] Flume PA, Liou TG, Borowitz DS, Li H, Yen K, Ordonez CL, et al. Ivacaftor in subjects with cystic fibrosis who are homozygous for the F508del-CFTR mutation. Chest 2012;142(3):718–24.

[49] Accurso FJ, Rowe SM, Clancy JP, Boyle MP, Dunitz PR, Sagel SD, et al. Effect of VX-770 in persons with cystic fibrosis and the G551D-CFTR mutation. NEJM 2010;363(21):1991–2003.

[50] Ramsey BW, Davies J, McElvaney NG, Tullis E, Bell SC, Drevinek P, et al. A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. NEJM 2011;365(18):1663–72.

[51] Heidelberger C, Chaudhuri NK, Danneberg P, Mooren D, Griesbach L, Dushinsky R, et al. Fluorinated pyrimidines: a new class of tumor inhibitory compounds. Nature 1957;179:663–6.

[52] Horie N, Aiba H, Ogura K, Hojo H, Takeishi K. Functional analysis and DNA polymorphism of the tandemly repeated sequences in the 59-terminal regulatory region of the human gene for thymidylate synthase. Cell Struct Funct 1995;20:191–7.

[53] Pullarkat ST, Stoelmacher J, Ghaderi V, Xiong YP, Ingles SA, Sherrod A, et al. Thymidylate synthase gene polymorphism determines response and toxicity to 5-FU chemotherapy. Pharmacogenomics J 2001;17(1):65–70.

[54] Carson PE, Flanagan CL, Ickes CE, Alving AS. Enzymatic deficiency in primaquine-sensitive erythrocytes. Science 1956;124:484–5.

[55] Mason PJ, Bautista JM, Gilsanz F. G6PD deficiency: the genotype-phenotype association. Blood Rev 2007;21(5):267–83.

[56] Youngster I, Arcavi L, Schechmaster R, Akayzen Y, Popliski H, Shimonov J, et al. Medications and glucose-6-phosphate dehydrogenase deficiency. Drug Safe 2010;33(9):713–26.

[57] Blanco-Molina A. Oral contraception in women with mild thrombophilia: what have we learned recently? Thromb Res 2012;130(Suppl 1):S16–8.

[58] Bertina RM, Koeleman BPC, Koster R, Rosendaal FR, Dirven RJ, de Ronde H, et al. Mutation in blood coagulation factor V associated with resistance to activate protein C. Nature 1994;369:64–7.

[59] Dayan N, Holcroft CA, Tagalakis V. The risk of venous thrombosis, including cerebral vein thrombosis, among women with thrombophilia and oral contraceptive use: a meta-analysis. Clin Appl Thromb Hemost 2011;17(6):E141–52.

[60] Poort SW, Rosendaal FR, Reitsma PH, Bertina RM. A common genetic variation in the 3' untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. Blood 1996;88:3698–703.

[61] Martinelli I, TaioloBucciarelli P, Akhavan S, Mannucci PM. Interaction between the G20210A mutation of the prothrombin gene and oral contraceptive use in deep vein thrombosis. Arterioscler Thromb Vasc Biol 1999;19(3):700–3.

[62] Lazarou J, Pomeranz BH, Corey PN. Incidence of ADRs in hospitalised patients. JAMA 1998;279:1200–5.

[63] Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TR, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 2004;329(7456):15–19.

[64] Martyrs C. Adverse reactions to drugs in general practice. BMJ 1979;2:1194–7.

[65] McCarthy TV, Quane KA, Lynch PJ. Ryanodine receptor mutations in malignant hyperthermia and central core disease. Hum Mutat 2000;15(5):410–7.

[66] Robinson R, Carpenter D, Shaw MA, Halsall J, Hopkins P. Mutations in RYR1 in malignant hyperthermia and central core disease. Hum Mutat 2006;27(10):977–89.

[67] Denbrough M. Malignant hyperthermia. Lancet 1998;352(9132):1131–6.

[68] Treves S, Anderson AA, Ducreux S, Divet A, Bleunven C, Grasso C, et al. Ryanodine receptor 1 mutations, dysregulation of calcium homeostasis and neuromuscular disorders. Neuromusc Disord 2005;15(9):577–87.

This page is intentionally left blank

Chapter 19

MicroRNA: Regulation of P450 and Pharmacogenetics

Miki Nakajima and Tsuyoshi Yokoi
Drug Metabolism and Toxicology, Faculty of Pharmaceutical Sciences, Kanazawa University, Japan

Chapter Outline

19.1 Overview 385
19.2 miRNA Biogenesis 385
19.3 miRNA Function 386
19.4 miRNA Target Identification 386
19.5 Regulation of Drug-Metabolizing P450s 389
  19.5.1 CYP2C8 389
  19.5.2 CYP2E1 389
  19.5.3 CYP3A4 390
19.6 Regulation of Cancer-Related P450s 391
  19.6.1 CYP1A1 391
  19.6.2 CYP1B1 392
  19.6.3 CYP2A 392
  19.6.4 CYP2J2 393
  19.6.5 CYP19A1 393
  19.6.6 CYP24A1 394
19.7 Modulation of miRNA Expression and its Toxicological Outcome 394
19.8 miRNA Pharmacogenetics Associated with Disease or Drug Response 395
19.9 miRNAs as Biomarkers and Therapeutic Targets 397
19.10 miRNAs and Drug Response 397
19.11 Conclusions 397
Glossary 398
Abbreviations 398
References 398

19.1 OVERVIEW

MicroRNAs (miRNAs) are short (~22-nucleotide), endogenous, noncoding RNAs that mediate gene silencing through translational repression or mRNA degradation [1]. Importantly, miRNAs are now recognized as critical regulators of gene expression. They regulate every cellular process and a large number of physiological processes [2]. Because the role of miRNAs is to fine-tune gene expression, thereby forming intricate networks, the dysregulation of miRNA function may lead to disease.

Cytochrome P450s (P450, CYP), a superfamily of hemoproteins, are the main enzymes in phase I drug metabolism. A major function of P450s is to convert nonpolar xenobiotics, including pharmaceuticals, carcinogens, and toxicants, as well as endogenous substrates such as steroids and fatty acids, into polar compounds. In addition to transcriptional regulation by transcription factors and nuclear receptors, it has been shown that P450s are directly or indirectly regulated by miRNAs. The latter mechanism impacts inter- and intraindividual variability in the metabolism of xeno/endobiotics. This chapter summarizes current knowledge of miRNA-regulating P450s and nuclear receptors, as well as the clinical significance of miRNA-related polymorphisms.

19.2 miRNA BIOGENESIS

miRNAs are first generated by RNA polymerase II as long primary transcripts called primary miRNAs (pri-miRNAs) (Figure 19.1). Each pri-miRNA forms a hairpin stem-loop structure that is cleaved by the RNase III enzyme Drosha in the nucleus, creating a precursor miRNA (pre-miRNA) that is 70–100 nucleotides in length. Then the pre-miRNA is exported into the cytoplasm by Exportin 5, processed into a miRNA duplex of ~22 nucleotides by Dicer or Argonaute 2 (Ago2), and then unwound into a single-stranded, mature miRNA. The mature miRNA, called the guide strand, is incorporated into the RNA-induced silencing complex (RISC), which is composed of Dicer, TAR RNA binding protein (TRBP) and Ago2; it then binds to its mRNA target [3]. The other strand, called miRNA* or the passenger strand, is degraded, although it is sometimes functional.

Thus, the biogenesis of miRNAs is regulated at the transcriptional and post-transcriptional levels. At the

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00019-0
© 2014 Elsevier Inc. All rights reserved.

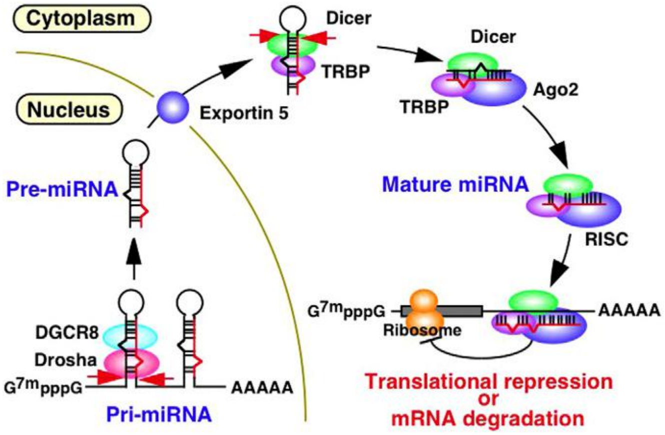


FIGURE 19.1 miRNA biogenesis and gene-silencing mechanisms. miRNA genes are transcribed by RNA polymerase II to generate pri-miRNA transcripts. The pri-miRNAs are processed by Drosha and DGCR8 in the nucleus to produce pre-miRNAs. Upon export by Exportin 5, Dicer processes the pre-miRNAs into miRNA duplexes. The duplexes are separated, and the guide strand is incorporated into miRISC, which recognizes the target mRNA and causes translational repression or mRNA degradation. *Source: Nakajima [93] with permission of the Showa University Journal of Pharmaceutical Sciences.*

transcriptional level, the production of miRNAs is controlled by mechanisms similar to those used to regulate protein-coding transcripts (e.g., transcription factors and epigenetics). At the post-transcriptional level, the biogenesis of miRNAs is regulated by miRNA processing machinery. New pathways for miRNA biogenesis have been recently reported and include the “mirtrons” pathway, in which a subset of miRNAs are derived from short hairpin introns, bypassing cleavage by Drosha, and the “simtrons” pathway, which generates splicing-independent mirtron-like miRNAs that require Drosha but not other canonical miRNA biogenesis components [4]. These recent findings suggest that regulation of miRNA expression is more complicated than expected. The control of miRNA biogenesis has been well characterized, but miRNA degradation is not fully understood. miRNAs have a long life span in general. However, the rapid turnover of miRNAs in a variety of specific cellular contexts has also been documented. Several ribonucleases that degrade miRNAs have been identified. The cellular contexts and biochemical mechanisms of miRNA decay are beginning to be uncovered [5].

To date, more than 2500 human miRNAs have been identified (miRBase, release 20, http://www.mirbase.org/). miRNAs are assigned sequential numerical identifiers [6]. Mature sequences are designated “miR,” whereas precursors are labeled “mir.” Abbreviated three- or four-letter prefixes are used to designate the species, such that identifiers take the form hsa-miR-101 (for miR-101 in *Homo sapiens*). The hsa-miR-101 in humans and the mmu-miR-101 in mice are orthologous. Paralogous miRNAs that differ at only one or two positions are given lettered suffixes (e.g., mmu-miR-10a and mmu-miR-10b). Distinct loci that give identical mature miRNAs have numbered suffixes (hsa-miR-125b-1 and hsa-miR-125b-2). The preferred method of naming guide/passenger (miRNA/miRNA*) strands designates miRNAs from 5′ or 3′ arms (e.g., miR-378-5p and miR-378-3p) [6].

19.3 miRNA FUNCTION

miRNAs function to negatively regulate gene expression via post-transcriptional mechanisms. They bind to their target mRNAs with imperfect pairing, usually in the 3′-untranslated region (UTR) or sometimes in the coding region to cause translational repression or mRNA degradation [3] (Figure 19.2). The mRNA sequence to which a miRNA binds is termed the miRNA recognition element (MRE). Gene silencing by miRNAs involves multiple mechanisms, including blocking of translation initiation, inhibition of ribosome elongation, degradation of mRNA by deadenylation or decapping, and relocalization of miRNAs associated with RISC to P-bodies, which are subcellular compartments enriched with mRNA degradation and translational repression factors [7]. miRNAs represent only 0.5–3% of the genome, but have been predicted to regulate approximately 60% of all human genes [8].

19.4 miRNA TARGET IDENTIFICATION

miRNAs bind to their targets by partial complementarity over a short sequence. During binding, nucleotides 2–8 at the 5′ end of the miRNA, called the seed sequence, (Figure 19.2), are critical [9], but perfect seed pairing alone may be insufficient. Sequences outside of the seed sequence, and even sequences outside of the MRE, can compensate or impact target regulation. Even sequences with high complementarity might not be the real targets of a miRNA. Therefore, the identification of miRNA(s) for a target gene is a challenging task. Multiple free databases (Table 19.1) are available to

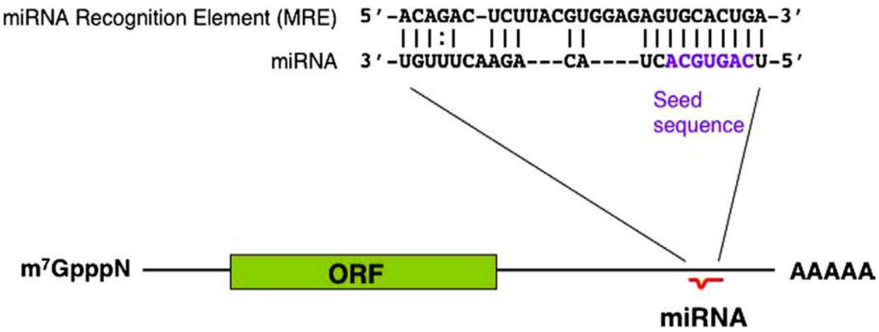

miRNA Recognition Element (MRE)
5'-ACAGAC-UCUUACGUGGAGAGUGCACUGA-3'
miRNA
3'-UGUUUCAAGA---CA----UCACGUGACU-5'
Seed sequence
m⁷GpppN
ORF
AAAAA
miRNA

FIGURE 19.2 **Binding of miRNAs to their target mRNAs.** miRNAs bind to their target mRNAs with imperfect pairing mainly within the 3′-UTR in mammals. The mRNA sequence to which a miRNA binds is termed the microRNA recognition element (MRE). During binding, nucleotides 2–8 at the 5′-end of the miRNA, called the seed sequence, are critical, but perfect seed pairing alone may be insufficient. Sequences outside of the seed sequence, and even those outside of the MRE, can compensate or impact the regulation of the target. ORF—open reading frame.

TABLE 19.1 Computational Tools for miRNA Target Prediction

Tool	Website
TargetScan	http://www.targetscan.org/
MicroCosm	http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/
PicTar	http://pictar.mdc-berlin.de/
DIANA MicroT	http://diana.cslab.ece.ntua.gr/microT/
TargetRank	http://genes.mit.edu/targetrank/
microRNA.org	http://www.microrna.org/microrna/home.do
PITA	http://genie.weizmann.ac.il/pubs/mir07/index.html

predict miRNAs for target genes. These algorithms place variable weight on complementarity to the miRNA seed sequence, evolutionary conservation of the MRE in the target gene, free energy of the miRNA–mRNA duplex, and accessibility of the target site (see Table 19.2).

It is difficult to say which algorithm produces the most reliable and/or sensitive target predictions, and the use of multiple programs is recommended. In most cases, many types of miRNAs are predicted for a given mRNA, and they vary significantly depending on the program used. The false-positive rate of the predicted candidate targets of a given miRNA is thought to be 30–50% [10,11]. Therefore, experiments are necessary to confirm these predictions. The following points can be taken into consideration to select miRNAs that are to be experimentally verified: (1) the agreement of multiple algorithm results; (2) the conservation of the target site sequence among species (points 1 and 2 are neither a necessity nor a guarantee of function; (3) accessibility to the target, which seems to be a critical feature of miRNA target recognition; and (4) whether the miRNA is substantially expressed in the tissue where the target mRNA is expressed.

The overexpression or inhibition of a miRNA is an effective method for determining potential miRNA–mRNA interactions. After the transfection of a precursor miRNA or antisense oligonucleotide (AsO) into cells to overexpress or inhibit the miRNA, the protein and mRNA levels of the target are examined. Because a possible concern of the miRNA overexpression experiment is that the decrease in target expression might be artificial under nonphysiological conditions, experiments in which endogenous miRNA is inhibited with an AsO are valuable. To comprehensively determine the targets of a miRNA, it is useful to examine the changes in mRNA or protein expression by microarray or proteome analysis after miRNA overexpression or inhibition [12,13]. It should be noted that the targets identified by this analysis include secondary targets, the expression of which may change as a result of the expression changes of the primary targets. We should also note that different proteins have different turnover rates. This means that the levels of proteins with high turnover will change rapidly, whereas stable proteins will be affected at later time points. In addition, if one only examines the changes in the mRNA levels, some targets whose downregulation is caused by translational repression may be missed. Thus, changes in both mRNA and protein levels should be considered for comprehensive analysis.

To verify the interaction between a miRNA and its target mRNA, an inverse relationship between the protein (or mRNA) levels and the miRNA level is necessary—for example, in cancerous versus normal tissues or among individuals or cell lines. To examine whether the mRNA is a direct target of the miRNA, luciferase assays are useful. The reporter constructs containing the MRE of the target mRNA downstream of the reporter gene are co-transfected into cells with a precursor miRNA or AsO. Evidence of an interaction can be established when the reporter activity is significantly decreased or increased.

TABLE 19.2 Potential miRNAs that may Bind to Sequences Around SNPs in P450 3′-UTR

Gene	dbSNP	miRNA Target Destroyed	miRNA Target Created
CYP1A2	rs34002060	miR-373*, miR-616*	miR-373*, miR-616*
CYP2A6	rs8192733	☐	miR-1225-3p
	rs28399469	☐	miR-328, miR-1291
CYP2B6	rs28969419	miR-876-3p	miR-323-5p
	rs3211391	miR-220b	☐
	rs3211392	miR-363*	☐
	rs3211393	miR-363*	☐
	rs3211398	miR-1179	☐
	rs34031833	miR-625	☐
	rs28399501	miR-501	☐
	rs3211372	miR-491, miR-663	☐
	rs7246465	☐	miR-570
	rs12979270	☐	miR-625
	rs1042389	☐	miR-650
	rs3211399	☐	miR-199a*
	rs3211403	☐	miR-214, miR-922
	rs3211404	miR-214	miR-376a*
CYP2C9	rs9332240	☐	miR-577
	rs9332241	☐	miR-641
	rs9332242	miR-143	☐
	rs9332243	☐	miR-136
	rs28969379	miR-183*	☐
CYP2E1	rs2480256	☐	miR-570-3p
	rs2480257	miR-570-3p	☐
CYP3A4	rs34469568	☐	miR-522
	rs34141651	☐	miR-522
	rs28988603	miR-148b*	☐
CYP3A5	rs17161788	☐	miR-330
	rs15524	miR-500a-5p, miR-501-5p, miR-362-5p	☐
CYP3A7	rs10211	☐	miR-125a-5p

Note: Box symbol indicates no miRNA.
Source: Adapted from Ramamoorthy and Skaar [77] and Wei et al. [78].

Different miRNAs may act together to regulate a target mRNA; this was corroborated by *in vitro* experiments. However, it is still unclear if different miRNAs act synergistically *in vivo*. In addition, the mRNA may be regulated independently by these miRNAs in different tissues or at different developmental stages, and this has been supported by *in vitro* cell-based experiments. Most studies have been performed *in vitro* using luciferase reporter systems and miRNA over- or misexpression experiments in cells. However, *in vivo* experiments to test whether individual miRNA targets are truly endogenous under endogenous conditions have been exceedingly rare.

19.5 REGULATION OF DRUG-METABOLIZING P450s

Cytochrome P450s (P450s, CYPs) are important enzymes that catalyze the detoxification of xenobiotics such as drugs, environmental chemicals, and carcinogens. The mechanisms of transcriptional regulation by nuclear receptors and the genetic polymorphism of P450s have been well studied to understand the reason for inter- and intraindividual variability in drug responses and adverse reactions, as well as xenobiotics-related toxicity. It was recently discovered that miRNAs are involved in the post-transcriptional regulation of P450s, conferring variability and modulation in drug detoxification and the metabolic activation of chemicals.

19.5.1 CYP2C8

The human CYP2C subfamily, which consists of CYP2C8, CYP2C9, CYP2C18, and CYP2C19, is responsible for the metabolism of > 20% of all clinically prescribed drugs. CYP2C protein and/or mRNA expression is at its highest levels in the human liver, but is also detected at lower levels in extrahepatic tissues such as the kidney, lung, heart, endothelium, adrenal gland, mammary gland, and brain. Numerous studies have described the transcriptional upregulation of *CYP2C* genes by xenobiotics, such as phenobarbital, rifampicin, St. John's wort, and dexamethasone. The post-transcriptional regulation of CYP2C8 by a miRNA was recently reported [14]. Potential MREs for miR-103 and miR-107 were identified in the 3′-UTR of CYP2C8 by an online search using the miRBase Targets database and TargetScan.

miR-103 and miR-107 contain the same seed sequence. When the precursor for miR-103 or miR-107 was transfected into primary human hepatocytes, the CYP2C8 protein level decreased, whereas AsO increased it. Neither the precursor nor the AsO affected the CYP2C8 mRNA level, indicating the post-transcriptional effect of these miRNAs. A putative MRE was found in the 3′-UTRs of CYP2C9 and CYP2C19, with a two- and one-nucleotide mismatch, respectively. Overexpression of miR-103 or miR-107 resulted in a decrease in the CYP2C9 and CYP2C19 protein levels in human hepatocytes, although to a lesser degree than CYP2C8 (CYP2C8 > CYP2C19 > CYP2C9).

miR-103 and miR-107, which belong to the miR-15/miR-107 group, are widely expressed in many human tissues [15]. They are transcribed from the introns of the pantothenate kinase (PANK) genes, which are involved in the regulation of acetyl-CoA levels, cell stress, insulin sensitivity, lipid metabolism, and angiogenesis. It would be of interest to examine the conditions that affect the expression of the PANK genes and the miR-103/miR-107 family and thus alter the expression of CYP2C8 and CYP2C19 in human tissues.

19.5.2 CYP2E1

Human CYP2E1 catalyzes the metabolism of low-molecular-weight xenobiotics, including drugs (e.g., acetaminophen, isoniazid, bromobenzene), organic solvents (e.g., ethanol, acetone, carbon tetrachloride, chloroform, glycerol, hexane, toluene), and procarcinogens (e.g., *N*-nitrosodimethylamine, *N*-nitrosopyrrolidine). CYP2E1 is induced by its own substrates, including isoniazid, ethanol, and acetone, through post-transcriptional or post-translational mechanisms. It is the most abundant isoform among all P450s in the human liver (56% of the total P450) at the mRNA level, followed by CYP2C9, CYP2C8, and CYP3A4 (8–11% of the total P450) [16], whereas it is the fourth most abundant isoform (approximately 7% of the total P450) at the protein level after CYP3A (30% of the total P450), CYP2C (20% of the total P450), and CYP1A2 (approximately 13% of the total P450).

It was reported that the CYP2E1 mRNA levels in 15 human liver samples were not positively correlated with chlorzoxazone 6-hydroxylase activities, which are probe activities of CYP2E1 [17]. Collectively, post-transcriptional regulation can be suggested not only for the inducible but also for the constitutive expression of CYP2E1 in the human liver.

The possibility that CYP2E1 is regulated by miRNA(s) has been studied [18]. In silico analysis identified a potential recognition element for miR-378 in the 3′-UTR of human CYP2E1 mRNA (Figure 19.3). Luciferase assays confirmed that miR-378 functionally recognizes the MRE. When the precursor miR-378 was transfected into cells expressing human CYP2E1 including its 3′-UTR, the CYP2E1 protein level and chlorzoxazone 6-hydroxylase activity were significantly decreased, but this result was not found in cells expressing CYP2E1 excluding its 3′-UTR. Interestingly, in both cell lines, the CYP2E1 mRNA levels were decreased by the overexpression of miR-378, although miR-378 did not affect the stability of the CYP2E1 mRNA. In a panel of 25 human liver tissues, no positive correlation was observed between the CYP2E1 protein level and the CYP2E1 mRNA level (Figure 19.3), supporting the

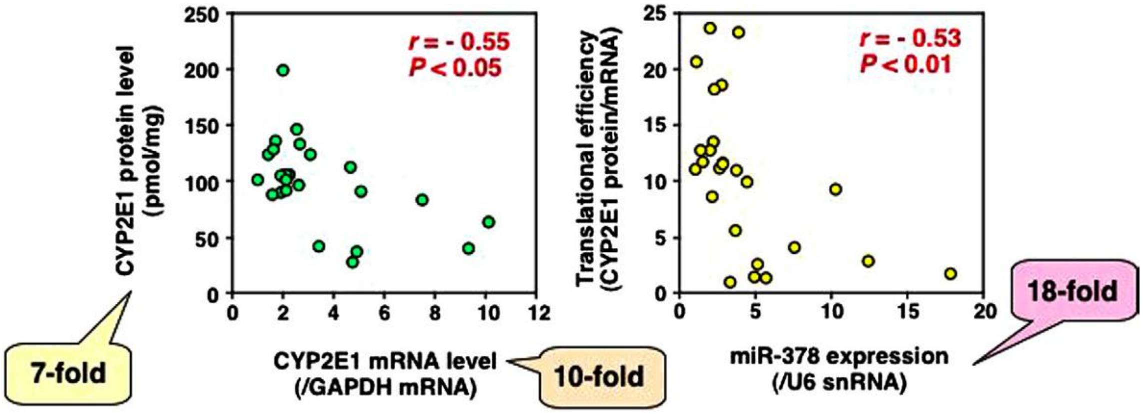


FIGURE 19.3 **Regulation of human CYP2E1 by miR-378.** miR-378 binds to the MRE in the 3′-UTR of CYP2E1 to repress its translation. In a panel of 25 human livers, no positive correlation was observed between the CYP2E1 protein level and the CYP2E1 mRNA level, whereas an inverse correlation was observed between CYP2E1 translational efficiency and the miR-378 level, indicating the impact of miR-378 on interindividual variability of human liver CYP2E1 expression. *Source: From Mohri et al.* [18] *with permission of Elsevier.*

post-transcriptional repression of CYP2E1. The miR-378 levels were inversely correlated with the CYP2E1 protein levels and the translational efficiency (protein/mRNA expression ratio) of human CYP2E1. This study provided new insight into the mechanism of the post-transcriptional regulation of CYP2E1. Because the 3′-UTR of CYP2E1 is poorly conserved among humans, rats, and mice, its regulation by miR-378 should be specific to humans.

The gene encoding miR-378 is within intron 1 of the *peroxisome proliferator-activated receptor-γ co-activator 1β* (*PGC1β*) gene on human chromosome 5q33.1 (Figure 19.4). This means that the expression of miR-378 should parallel that of *PGC1β*. *PGC1β* is a transcriptional coactivator that regulates metabolism and mitochondrial biogenesis through the stimulation of nuclear hormone receptors and other transcription factors. It has been reported that the expression of *PGC1β* is downregulated in diabetes and obesity but upregulated by insulin treatment [19,20]. In contrast, the expression of CYP2E1 is upregulated in diabetes and obesity but is downregulated by insulin treatment [21–23].

miR-378 may underlie the opposing results for *PGC1β* and CYP2E1. It was recently demonstrated that mice genetically lacking miR-378 and miR-378* are resistant to high fat diet-induced obesity, and exhibit enhanced mitochondrial fatty acid metabolism and elevated oxidative capacity in insulin target tissues [24]. Carnitine *O*-acetyltransferase (a mitochondrial enzyme involved in fatty acid metabolism) and MED13 (a component of the Mediator complex that controls nuclear hormone receptor activity), which are implicated in metabolic syndrome and obesity, were elevated in the livers of miR-378/378* knockout mice. Thus, miR-378 and miR-378* seem to be integral components of a regulatory circuit that functions under conditions of metabolic stress to control systemic energy homeostasis and the overall oxidative capacity of insulin target tissues.

19.5.3 CYP3A4

CYP3A4 is the most important P450 enzyme, as it facilitates the metabolism of a wide range of structurally different xenobiotics, including more than 50% of all clinically relevant drugs. There are large (∼50-fold) interindividual differences in the level of CYP3A4 expression in the human

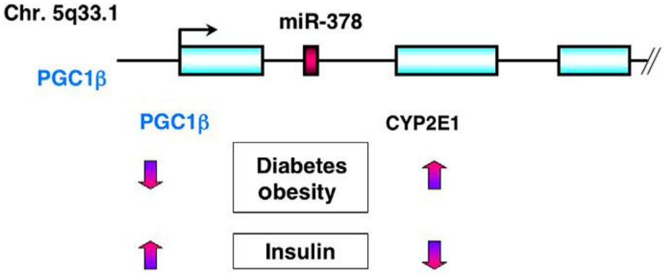


FIGURE 19.4 **Reciprocal relationship of miR-378 and CYP2E1 expression levels.** The gene encoding miR-378 is within intron 1 of the *PGC1β* gene on human chromosome 5q33.1. This means that the expression of miR-378 should parallel that of *PGC1β*, which is known as a regulator of hepatic lipid synthesis and lipoprotein production. It has been reported that the expression of *PGC1β* is downregulated in diabetes and obesity but upregulated by insulin treatment. In contrast, the expression of CYP2E1 is upregulated in diabetes and obesity but downregulated by insulin treatment. miR–378-dependent regulation may underlie the opposing changes in *PGC1β* and CYP2E1 expression.

liver and at least 6-fold interindividual variations in the activity of CYP3A4 in most populations. This variability cannot be explained solely by genetic polymorphisms [25]. CYP3A4 expression is predominantly regulated at the transcriptional level, mainly by pregnane X receptor (PXR). This idea is supported by the fact that the PXR protein level was significantly correlated with the CYP3A4 mRNA and protein levels in a panel of 25 human livers [26]. Because the CYP3A4 protein level was positively correlated with the CYP3A4 mRNA level, even if miRNA(s) are involved in the regulation of CYP3A4, their contribution may be masked by the impact of transcriptional regulation. In contrast, the PXR protein level was not correlated with the PXR mRNA level, suggesting its post-transcriptional regulation.

In this study, it was clearly demonstrated that miR-148a binds to an MRE in the 3′-UTR of human PXR mRNA to repress its translation (Figure 19.5). Because the miR-148a level in the liver panel was inversely correlated with the translational efficiency of PXR (PXR protein/PXR–mRNA ratio), miR-148a-dependent regulation can be said to be functional in the human liver, affecting the constitutive and inducible expression of CYP3A4. Interestingly, the induction of CYP2B6 (2-fold) and multidrug resistance 1 (MDR1)/P-glycoprotein (5-fold) mRNA by rifampicin, a typical ligand of human PXR, in LS180 cells was attenuated by the overexpression of miR-148a [26]. Accordingly, the miR-148a-dependent regulation of PXR appeared to affect downstream genes in common.

Furthermore, it was reported that the CYP3A4 protein level in LS180 and human pancreatic cancer-derived PANC1 cells was decreased by the overexpression of miR-27b and was accompanied by a decrease in CYP3A4 mRNA levels [27]. This finding is based only on an overexpression study. To demonstrate the importance of the miR-27b-dependent regulation of CYP3A4, an inhibition study of endogenous miR-27b, as well as a correlation analysis between the miRNA level and the target level, would be worthwhile.

19.6 REGULATION OF CANCER-RELATED P450s

Accumulating evidence shows that miRNA expression in cancer tissues is quite different from that in normal tissues. While certain miRNAs have been found to serve as oncogenes or tumor suppressor genes, these miRNAs and their contribution to the long process of cancer initiation or progression remain largely unknown. P450s catalyze the metabolic activation of a variety of procarcinogens and promutagens as well as the metabolism of endogenous compounds, including steroids, that are associated with cancer. Thus, they are associated with the initiation or progression of cancer. Some cancer-associated P450s are regulated by miRNAs, and this regulation seems to be an underlying mechanism of cancer initiation/progression.

19.6.1 CYP1A1

CYP1A1 is not constitutively expressed in liver, but it is inducible in the liver and many extrahepatic tissues, including the lung. The induction of CYP1A1 by polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls, and halogenated dioxins, which can be found in charcoal-grilled foods or the atmosphere, is dependent on the activation of the aryl hydrocarbon receptor (AhR) and

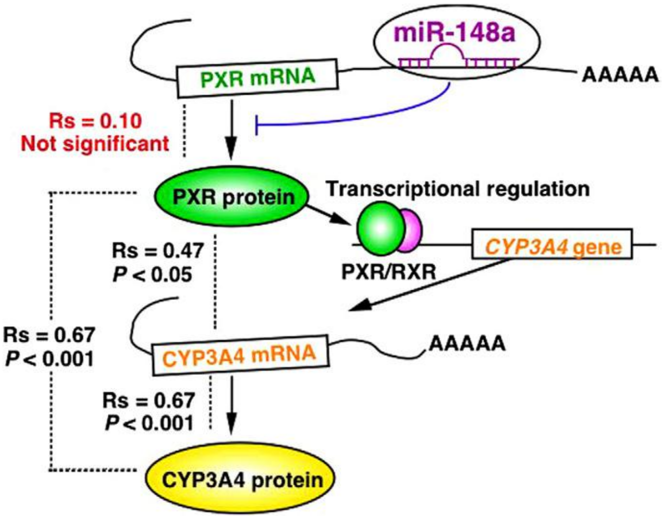


FIGURE 19.5 miR-148–dependent regulation of human PXR, affecting CYP3A4 expression in the liver. miR-148 binds to an MRE in the 3′-UTR of PXR to repress its translation. In a panel of 25 human livers, a large interindividual variability (100-fold) was found in the miR-148 expression level. An inverse correlation was observed between the translational efficiency of PXR and the expression of miR-148a. Because the PXR protein level was significantly correlated with CYP3A4 mRNA and protein levels, the miR-148a-dependent regulation of PXR is implicated as an underlying mechanism of the large interindividual variability in CYP3A4 expression.

therefore transactivation. CYP1A1 catalyzes the bioactivation of procarcinogens and proteratogens to active metabolites that bind to DNA.

It was found that miR-892a negatively regulates human CYP1A1 expression [28]. Luciferase assays revealed that an MRE in the 3′-UTR that shows a perfect match with miR-892a is functional. Overexpression or inhibition of miR-892a in MCF-7 cells resulted in a decrease or increase in CYP1A1 expression, respectively. Interestingly, benzo(*a*)pyrene, a potent inducer of CYP1A1 transcription, decreased the expression of miR-892a. The miR-892a-induced CYP1A1 repression inhibited the benzo(*a*)pyrene-mediated decrease in cell viability.

Recently, Oda et al. [29] reported that the aryl hydrocarbon receptor nuclear translocator (ARNT), a heterodimer partner of AhR, or hypoxia inducible factor 1α (HIF-1α) in the human liver, is regulated by miR-24. Overexpression of miR-24 in HuH-7 and HepG2 cells significantly decreased the ARNT protein level, but not the ARNT mRNA level, indicating its translational repression. The miR-24-dependent downregulation of ARNT decreased the expression of its downstream genes, such as carbonic anhydrase IX and CYP1A1. Taken together, CYP1A1 is directly and indirectly regulated by miRNAs.

19.6.2 CYP1B1

CYP1B1 is expressed in ovarian, uterine, and breast tissues and plays a key role in estrogen metabolism. 4-Hydroxyestradiol, a catechol-type metabolite of estradiol that is produced by CYP1B1, generates free radicals from reductive-oxidative cycling with its corresponding semiquinone and quinone forms, which causes DNA damage and contributes to estrogen-dependent cancers, such as breast and endometrial cancers. In addition, CYP1B1 catalyzes the metabolic activation of a variety of procarcinogens and promutagens, including PAHs and aryl amines. It was demonstrated that the CYP1B1 protein level is higher in various types of cancer compared with normal tissue, whereas there is no difference in the CYP1B1 mRNA levels between cancerous and normal tissues. Although there was no direct evidence to show a lack of association between the mRNA and protein levels of CYP1B1 in a panel of human tissues, these phenomena are reminiscent of post-transcriptional regulation.

Tsuchiya et al. [30] found that human CYP1B1 is post-transcriptionally regulated by miR-27b via translational repression. This was the first example of a P450 enzyme regulated by miRNA. The 3′-UTR of human CYP1B1 is uniquely long (3.1 kb), and an MRE for miR-27b was found in a highly conserved region at nucleotides 4358–4381. A luciferase assay clearly demonstrated that the MRE in the 3′-UTR, which is located in a highly conserved region between species, was functional in miR-27b binding. Interestingly, the expression level of miR-27b was lower in breast cancer tissues than in adjacent normal tissues. A significant inverse association was observed in the expression levels of miR-27b and the CYP1B1 protein between breast cancer and normal tissues. The decreased expression of miR-27b may be one of the causes of the high expression of the CYP1B1 protein in cancer tissues (Figure 19.6).

Human CYP1B1 is unique in that it has a long and complex 5′-UTR (373 nucleotides). After the study by Tsuchiya et al. [30], it was reported that the 5′-UTR of CYP1B1 interacted with the MRE in the 3′-UTR to almost completely inhibit translation [31]. It is therefore likely that the absence of the CYP1B1 protein in normal cells is due to the high degree of translational control exerted through its 5′- and 3′-UTR.

Human CYP1B1 is also regulated by estrogen receptor α (ERα) at the transcriptional level [32]. Human ERα is regulated by miR-206, and its activation results in decreased expression of miR-206, showing mutually inhibitory regulation [33]. miR-221, miR-222 [34], and miR-22 [35] were also reported to regulate human ERα expression. The effects of the miRNA-dependent regulation of ERα on CYP1B1 expression remain to be studied.

19.6.3 CYP2A

Human CYP2A13 is the major P450 enzyme in the respiratory tract. It catalyzes the α-hydroxylation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and its metabolites, activating and rendering them mutagenic [36]. Therefore, CYP2A13 is critically involved in eliciting the biological effects of NNK. The orthologs of CYP2A13 in rodents are CYP2A3 in rats, and Cyp2a4 and Cyp2a5 in mice.

It was reported that the chronic administration of NNK to F344 rats for up to 20 weeks resulted in decreased expression of several miRNAs, including miR-126* and miR-34 in the lung [37]. A luciferase assay revealed that an MRE in the 3′-UTR of CYP2A3 was functionally recognized by miR-126*. In the lungs of the NNK-treated rats, the CYP2A3 mRNA and protein levels were significantly increased compared with control rats. This is a possible mechanism that may potentiate NNK genotoxicity. Because the MRE sequence for miR-126* is highly conserved among species, and because miR-126* has been suggested as a potential miRNA for human CYP2A13 by in silico programs, it is feasible that CYP2A13 is also regulated by miR-126*.

Another miRNA with clearly interesting implications is miR-34 [37], which is known to inhibit the expression of a large number of genes involved in DNA damage response, cell cycle progression, and apoptosis. The decrease in miR-34 expression after NNK exposure upregulated many genes that control important processes, such as cell proliferation and cell death, which most likely contribute to lung cancer formation [38].

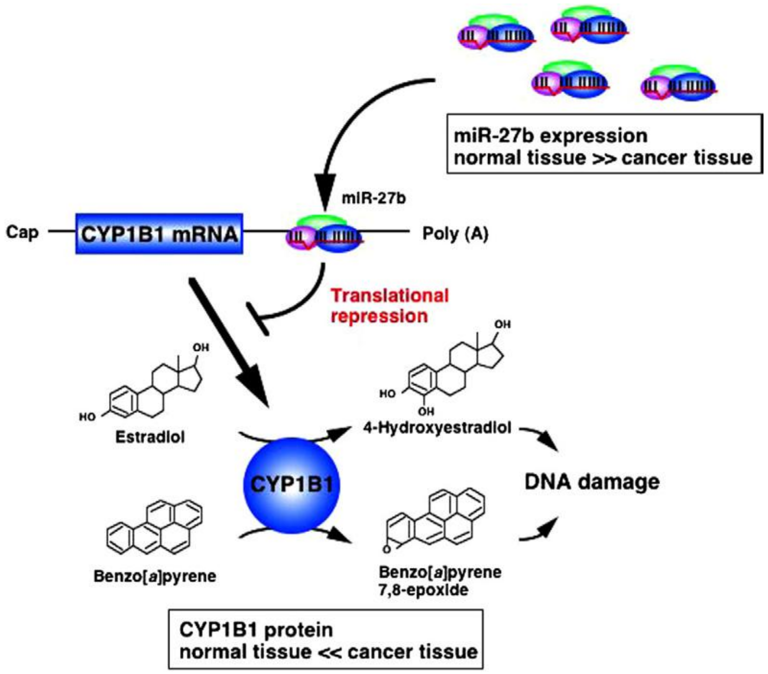


FIGURE 19.6 **Regulation of human CYP1B1 by miR-27b and dysregulation in cancer tissues.** miR-27b binds to a MRE in the 3′-UTR of CYP1B1 to repress its translation. The expression level of miR-27b in cancer tissue is lower than in normal tissue, which could be a cause of the high expression of CYP1B1 in cancer tissue.

19.6.4 CYP2J2

Human CYP2J2 is primarily expressed in the heart and vessel endothelium and is also found in a variety of tissues, including liver, lung, kidney, and gastrointestinal tissues [39]. This enzyme catalyzes the epoxidation of arachidonic acid into four regioisomers of *cis*-epoxyeicosatrienoic acid (5,6-EET, 8,9-EET, 11,12-EET, and 14,15-EET) [40]. EETs have diverse biological effects in the cardiovascular system, and they play important roles in processes related to cancer cell behavior and tumor pathogenesis. Exogenous EETs markedly accelerate the proliferation and metastasis of cancer cells *in vitro* and *in vivo*.

The association of CYP2J2 with cancer is supported by the fact that CYP2J2 is highly expressed in human tumors but not in adjacent normal tissues or nontumoral human cell lines [41]. It was demonstrated that the protein level and enzyme activity of human CYP2J2 in several cancer cell lines, including HeLa and MDA-MB-435, were significantly decreased by the overexpression of let-7b [40], a known tumor suppressor miRNA. A luciferase assay revealed that one out of seven potential MREs for let-7b in the 3′-UTR of CYP2J2 was functional for its regulation. Interestingly, an inverse correlation was demonstrated between the let-7b expression level and the CYP2J2 protein level in 18 sets of lung squamous tumor cancer and paired adjacent nontumor tissues [40]. Let-7b significantly inhibited the tumor phenotypes of tumor-xenografted mice by targeting CYP2J2. Thus, the repression of CYP2J2 was added to the list of functions for let-7b as a tumor suppressor miRNA.

19.6.5 CYP19A1

CYP19A1 (aromatase) plays a critical role in estrogen biosynthesis. To date, human CYP19A1 has been reported to be regulated by four different miRNAs: let-7f, miR-378, miR-181a, and miR-98. Xu et al. [42] demonstrated that miR-378 suppresses CYP19A1 protein expression and estradiol production in ovarian granulosa cells. A luciferase assay revealed that CYP19A1 is a direct target of miR-378. Panda et al. [43] reported that miR-98, whose aberrant expression is associated with the endometrial transition from a normal to a cancerous state, suppresses CYP19A1 mRNA and protein expression through direct interaction with its 3′-UTR in the human endometrial adenocarcinoma cell line Ishikawa. Shibahara et al. [44] reported that CYP19A1 mRNA is significantly increased in MCF-7 cells by co-culture with primary stromal cells from human breast cancer tissues, and is downregulated by subsequent treatment with letrozole, an aromatase inhibitor. In letrozole-treated MCF-7 cells, 13 miRNAs were upregulated, including let-7f, a known tumor suppressor miRNA. Multiple target-prediction algorithms suggested let-7f as a regulator of CYP19A1, and a luciferase assay revealed that an MRE in the 3′-UTR of CYP19A1 is functional, indicating that CYP19A1 is a direct target of let-7f. These results indicate that aromatase inhibitors may exert tumor-suppressing effects on breast cancer cells by suppressing CYP19A1 expression via restoration of let-7f. An inverse correlation between let-7 expression and the

CYP19A1 mRNA and protein levels in 11 breast cancer tissues supported the clinical significance of this regulatory mechanism.

19.6.6 CYP24A1

Human CYP24A1 is an essential enzyme for the inactivation of 1,25-dihydroxyvitamin D3 (calcitriol), a biologically active metabolite of vitamin D3. Calcitriol is typically considered to be a regulator of calcium homeostasis, but has received much attention for its antitumor activity. Most of the biological effects of calcitriol are elicited by its binding to the vitamin D receptor (VDR) [45]. It was demonstrated that the expression of CYP24A1 [46] and VDR [47,48] is higher in cancer tissues than in normal tissues at the protein level, but is not necessarily accompanied by an increase in the mRNA level. This suggests the involvement of post-transcriptional regulation.

Human CYP24A1 [49] and VDR [50] are both regulated by miR-125b (Figure 19.7). Because it is transcriptionally regulated by VDR, its repression by miR-125b should occur via both direct and indirect mechanisms. The miR-125b level was found to be lower in breast cancer tissues than in normal tissues, indicating a possible cause of high CYP24 expression in cancer [49]. An increase in VDR in cancer tissues would augment the antitumor effects of calcitriol, whereas an increase of CYP24A1 would attenuate them. Because the antiproliferative effects of calcitriol in MCF-7 cells were significantly abolished by the overexpression of miR-125b (Figure 19.8), it was suggested that miR-125b has a larger impact on the regulation of VDR than on the regulation of CYP24A1 in this cell system.

miR-125b is considered a tumor suppressor gene [51]. It inhibits the proliferation of human breast cancer cells [51], hepatocellular carcinoma cells [52], and thyroid carcinoma cells [53]. In contrast, its inhibition decreases the growth of human prostate cancer cells. miR-125b expression shows differential changes in human tumors—for example, it is downregulated in breast, ovarian, and bladder cancer but is upregulated in pancreatic and stomach cancer [54]. Therefore, the functional effects of miR-125b may differ between cancer types.

19.7 MODULATION OF miRNA EXPRESSION AND ITS TOXICOLOGICAL OUTCOME

Several lines of evidence indicate that miRNA expression is changed by exposure to chemicals, drugs, or hormones. It is conceivable that the up- or downregulated miRNAs may alter the expression of numerous target mRNAs, thus causing toxicological phenotype changes. It was reported that administering 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) to mice and rats caused changes in the levels of 45 and 17 miRNAs, respectively, in the liver [55]. Mouse and rat hepatoma cells in culture also exhibited changes in their miRNA levels in response to TCDD treatment [55]. Although TCDD is well known to upregulate the expression of numerous genes via AhR, it can also downregulate the expression of some genes. While researchers surmised that miRNAs might be responsible for mRNA downregulation in the TCDD/AhR pathway, it seems that mRNA downregulation is unlikely to be mediated by miRNAs and that miRNAs do not play a significant role in TCDD toxicity in the adult rodent liver.

It was demonstrated by *in vivo* and *in vitro* studies that WY-14,643, a potent peroxisome proliferator-activated receptor α (PPARα) agonist decreases the expression in mice

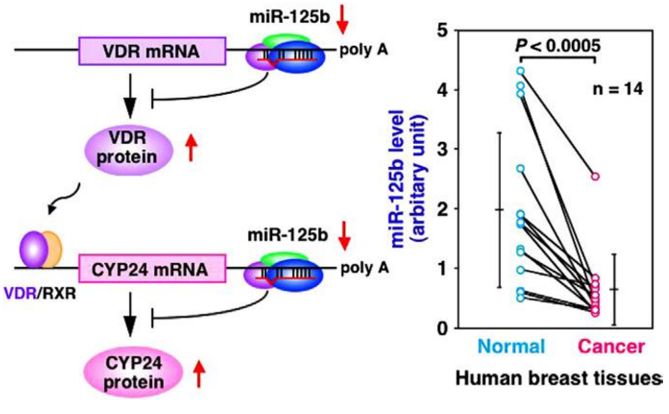


FIGURE 19.7 **Regulation of human CYP24A1 and VDR by miR-125b.** Human CYP24A1 and VDR are both regulated by miR-125b through MREs in their 3′-UTRs. Because CYP24A1 is transcriptionally regulated by VDR, its repression by miR-125b occurs via both direct and indirect mechanisms. The expression level of miR-125b in breast cancer tissue is lower than in normal tissue, which could be a cause of the high expression of CYP24A1 and VDR in cancer tissue. *Source: Nakajima* [93] *and Komagata et al.* [49] *with permission of the Showa University* Journal of Pharmaceutical Sciences *and the American Society for Pharmacology and Experimental Therapeutics, respectively.*

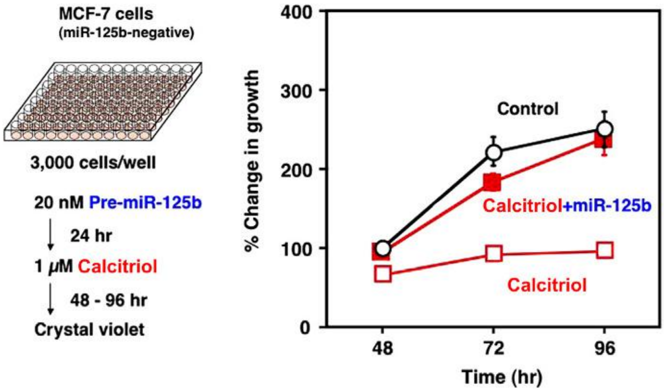


FIGURE 19.8 **Effects of miR-125b overexpression on the anti-proliferative effects of calcitriol.** The precursors for miR-125b or control were transfected into MCF-7 cells. After 24 h, the cells were treated with calcitriol for 48–96 h; crystal violet assays were then performed. The overexpression of miR-125b prominently abolished the antiproliferative effects of calcitriol. miR-125b, which regulates VDR, has a great impact on calcitriol's antiproliferative effects. In cancer cells, the downregulation of miR-125b resulted in the augmentation of calcitrol's antitumor effects.

of let-7c, which is an important miRNA in cell growth [56]. Because let-7c negatively regulates c-*myc* expression, the PPARα-mediated induction of c-*myc* via let-7c subsequently increases expression of the oncogenic miR-17-92 cluster. Thus, the let-7c signaling cascade seems to be an underlying mechanism of PPARα agonist–induced liver proliferation.

Several toxicogenomic studies have been performed to evaluate the responses or roles of miRNAs in toxicity and to establish biomarker(s) for toxicological risk assessment. Some chemicals that cause liver injury have been examined. Single administration of acetaminophen or carbon tetrachloride to rats causes changes in hepatic miRNA expression [57]. In a search for miRNAs that are decreased by two hepatotoxicants, miRNA-298 and miR-370 were identified. These are predicted to regulate thioredoxin reductase 3, an enzyme that responds to oxidative stress. Because changes in expression of these miRNAs were observed in the early phase of toxicity, they may be related to the induction of hepatotoxicity.

Chronic exposure of rats to tamoxifen, a potent hepatocarcinogen, for 24 weeks leads to substantial changes in liver miRNA expression (20 upregulated and 13 downregulated miRNAs) [58]. A significant upregulation of oncogenic miRNAs, such as the miR-17-92 cluster, miR-106a, and miR-34, was observed. In addition, the downregulation of miR-152 and miR-195, which are known to be downregulated in solid tumors [59], was found. The corresponding changes in the expression of the proteins targeted by these miRNAs, which include important cell cycle regulators, chromatin modifiers, and expression regulators implicated in carcinogenesis, were observed. These miRNA changes occur prior to tumor formation and are not merely a consequence of a transformed state.

Recently, Zhang et al. [60] demonstrated that estrogen inhibits CCl_4-induced hepatic injury through the induction of hepatic miR-29 *in vivo* in mice. Previous investigations have demonstrated that female animals are generally more resistant to CCl_4-induced fibrosis than males, owing to the protective effect of estradiol. The members of the miR-29 family, miR-29a and miR-29b, are key regulators of collagen and are implicated in tissue fibrosis [61]. Both estradiol treatment and intravenous injection of a recombinant adenovirus expressing miR-29a/b markedly attenuates the expression of fibrotic markers in mouse livers during CCl_4 treatment. This supports the protective role of estradiol-induced miR-29 in liver injury [60]. The modulation of miRNA expression by chemicals, drugs, or hormones is involved in the occurrence or mitigation of toxicological responses.

19.8 miRNA PHARMACOGENETICS ASSOCIATED WITH DISEASE OR DRUG RESPONSE

Pharmacogenetics research has matured considerably over the past few decades. Single-nucleotide polymorphisms (SNPs), the most common human genetic variants, can affect either the expression or the activity of various enzymes and may therefore be associated with differences in physiological or pharmacological outcomes. Recent genome-wide analyses of human SNPs have revealed that many of them exist in miRNA binding sites in the 3′-UTR. Approximately 400 SNPs were found at verified target sites or predicted target sites, and approximately 250 SNPs potentially create novel target sites for miRNAs in humans [62]. Such gain- or loss-of-function mutations result in changes in gene expression.

Interestingly, many examples of the association of SNPs in miRNA target sites with diseases [63] have been reported. The T>G SNP in the 3′-UTR of the *KRAS* gene, which abrogates the binding of let-7, is associated with increased risk of non-small-cell lung cancer [64]. The

G>A SNP in the 3′-UTR of the human *Slit and Trk-like 1* gene, which is implicated in Tourette's syndrome and attention deficit hyperactivity disorder, was found to strengthen the miR-189 target site, resulting in the gene's more stringent regulation [65]. A 1166A>C SNP (rs5186) in the *angiotensin receptor 1* (*AGTR1*) gene, which is implicated in hypertension and cardiovascular disease, was shown to abrogate miR-155-mediated regulation, resulting in overexpression of the AGTR1 protein [66]. The SNP in the 3′-UTR of the human *fibroblast growth factor 20* (*FGF20*) gene, which is implicated in Parkinson disease, was found to abrogate miR-433-mediated regulation, resulting in the overexpression of the FGF20 protein and its target gene α-*synuclein*, which causes Parkinson's disease [67]. The G>A SNP (rs62625044) in the 3′-UTR of the human *serotonin receptor type 3 subunit (HTR3A)* gene, which shows a strong association with female irritable bowel syndrome, was shown to abrogate miR-510-mediated regulation, resulting in the overexpression of the receptor subunit [68].

Studies of SNPs in the 3′-UTR that change the miRNA recognition of genes related to drug response are still immature. The first evidence of this phenomenon was reported for the human *dihydrofolate reductase* (*DHFR*) gene [69]. The 829C>T SNP of the *DHFR* gene (rs5030762), which is located 14 bp downstream of a miR-24 binding site in the 3′-UTR, was demonstrated to cause the overexpression of *DFHR* and methotrexate resistance [69] by evading negative regulation by miR-24. This finding suggests that sequences outside of the miRNA binding site may also be important.

However, a recent study [70] reported that the 829C>T SNP was an artifact derived from the *DHFRP1* pseudogene, which has the corresponding sequence 829T. Originally, the 829C>T SNP was genotyped by the polymerase chain reaction–restriction fragment-length polymorphism (PCR-RFLP) using cDNA samples prepared from total RNA samples that were extracted from peripheral blood mononuclear cells and a primer pair set on exons [69,71]. The *DHFRP1* gene on chromosome 18 lacks introns and has an open reading frame that is virtually identical to the coding sequence of the normal *DHFR* gene (on chromosome 5) [72]. When PCR-RFLP was performed to genotype the 829C>T SNP using genomic DNA samples with an intronic sense primer, no samples with the 829T allele were found [70]. Thus, the presence of pseudogene(s) should receive special attention with respect to genotyping and quantification at the mRNA level. The relationship between the 829C>T SNP, the DFHR mRNA level, and the involvement of miR-24 should be carefully investigated in future studies.

Another example of a SNP that affects miRNA recognition is the 973C>T SNP (rs1042157) in the *SULT1A1* gene, which encodes a sulfotransferase isoform that catalyzes the metabolism of a wide variety of endo- and xenobiotics, including dietary and environmental procarcinogens and promutagens [73]. The T allele is associated with the increased affinity of miR-631 to *SULT1A1* mRNA, leading to the gene's decreased expression.

The importance and implications of genetic polymorphisms in genes encoding drug-metabolizing enzymes are largely recognized. To date, the effects of SNPs in the coding region or 5′-flanking region on enzyme activity or expression have been eagerly investigated, but SNPs within the 3′-UTR have been largely overlooked because they are unlikely to affect enzymatic properties or expression levels. With the knowledge that miRNAs bind to sequences in the 3′-UTR, we need to reconsider the significance of SNPs in this region. Much information on human P450 variants and their outcomes has been compiled in the human cytochrome P450 (CYP) allele nomenclature database (http://www.cypalleles.ki.se/).

There are many SNPs in the 3′-UTRs of P450s. Interestingly, it has been reported that polymorphisms in the 3′-UTRs of CYP2A6 [74,75] and CYP2E1 [76] are associated with altered expression or activity. It would be of interest to investigate the possibility that regulation by miRNAs might be involved in this phenomenon. Furthermore, using in silico analysis, several research groups have reported potential miRNAs that may bind to the sequences around SNPs in the 3′-UTRs of drug-metabolizing P450s, such as CYP1A2, CYP2A6, CYP2B6, CYP2C9, CYP2E1, CYP3A4, CYP3A5, and CYP3A7, thus destroying or creating miRNA target sites [77,78]. Experimental proof is required to ascertain whether the predicted miRNAs actually regulate these P450s and so cause SNP-related interindividual variability in their expression levels.

SNPs are present not only in mRNA sequences but also in mature miRNA sequences. SNPs in mature miRNA, pri-miRNA, and pre-miRNA could modify various biological processes by influencing target selection or miRNA processing. When a SNP exists in the seed sequence of mature miRNAs, it can affect miRNA–mRNA interactions. When a SNP exists in the pri-miRNA or pre-miRNAs sequence, it may affect the expression level of the corresponding mature miRNA [79,80] and lead to modifications in target gene expression levels. Compared to the considerable level of SNPs in protein-coding mRNA, the SNPs in pri-miRNA, pre-miRNA, and mature miRNA are relatively rare. It has been reported that only ~10% of human pre-miRNAs have SNPs and that <1% of mature miRNAs have SNPs in their functional seed sequences [62].

Any given miRNA may regulate hundreds of different targets at different spatial-temporal settings. Therefore, functional mutations in miRNAs as top-level regulators are unlikely to be tolerated and are removed by purifying selection [62]. Although variations in the miRNA seed region seem to be very rare and occur with low frequency, they have the potential to influence the expression of hundreds of genes and related pathways.

In recent years, SNPs in miRNAs that alter their processing were reported to be associated with cancer risk, tumor response to anticancer agents, time to progression in cancer patients, and ulcerative colitis incidence [81]. However, in most cases, the molecular mechanism by which these polymorphisms act is not yet understood. For example, SNPs in pre-miR-196a2 (rs11614913) and pre-miR-499 (rs3746444) are significantly associated with an increased risk of breast cancer [82]. A SNP in pre-miR-27a is associated with the development of gastric mucosal atrophy in Japanese men [83]; also, a SNP in pre-miR-146a (rs2910164), which reduces premature and mature miR-146a levels, leads to a less efficient inhibition of target genes involved in the Toll-like receptor and in cytokine-signaling pathways and are associated with an increased risk of papillary thyroid carcinoma [84]. Furthermore, SNPs in pri-miR-26a (rs7372209) and pri-miR-100 (rs1834306) are significantly associated with tumor response or time to progression in 61 metastatic colorectal cancer patients treated with 5-fluorouracil and irinotecan [85]. This is the first report to reveal a relationship between the clinical outcome of drug therapy and SNPs in miRNA genes.

An important topic for future research is the effect of SNPs in miRNA sequences on drug metabolism and pharmacokinetics. A polymorphism in an mRNA target site would be target-specific, whereas a polymorphism in a miRNA may affect the expression of multiple genes and have serious consequences. The pharmacogenetics of miRNAs represents an innovative field of research for predicting drug response and toxicity. Increased knowledge of miRNA-related polymorphisms may be useful in the development of more successful individualized medicine. In this regard, the following databases are useful for investigating miRNA-related polymorphisms:

- Patrocles (http://www.patrocles.org/Patrocles.htm)
- PolymiRTS (http://compbio.uthsc.edu/miRSNP/)
- MicroSNiPer (http://cbdb.nimh.nih.gov/microsniper/)
- miRNA SNiPer (http://www.integratomics-time.com/miRNA-SNiPer/)
- miRNASNP (http://www.bioguo.org/miRNASNP/)
- miRdSNP (http://mirdsnp.ccr.buffalo.edu/)

19.9 miRNAs AS BIOMARKERS AND THERAPEUTIC TARGETS

Because of the growing evidence that a variety of diseases are associated with the dysregulation of miRNAs, miRNAs are now considered to be new tools for diagnosis and therapy [86]. Especially since the discovery of circulating and extracellular miRNAs in 2008, the study of miRNAs in body fluids, such as serum, plasma, urine, and saliva has rapidly expanded. Numerous studies have shown that circulating or extracellular miRNAs can be implicated in pathogenic conditions, such as various cancers, heart disease, diabetes, and organ damage. It is recognized that the miRNA profile of body fluids reflects the disease state as well as the presence of organ damage and injury. For this reason, circulating miRNAs are receiving much attention as a novel class of disease biomarkers.

Although the precise role of extracellular miRNAs is still largely unknown, it has also been shown that subcellular vesicles containing miRNA can be taken up by cells, causing changes in cellular gene expression in pathological conditions. These findings suggest biological functions for extracellular miRNAs that have yet to be fully investigated.

The modulation of miRNA levels is a potential therapeutic target because miRNAs are differentially expressed in diseases and play critical roles in various biological pathways. Some pioneering pharmaceutical companies have initiated the development of therapeutic candidates using miRNA mimics or inhibitors for cancer, cardiovascular diseases, neurological disorders, and viral infections. There are several ongoing clinical trials [87,88]. Because an individual miRNA could regulate several genes and pathways simultaneously, miRNA modulation could produce a powerful outcome. However, attention must be paid to the possibility that miRNA manipulation causes unanticipated effects, as the targets of each miRNA have yet to be thoroughly clarified. Although there are some hurdles to be overcome, such as stability, appropriate *in vivo* delivery systems, and selectivity, targeting miRNAs for therapy is an emerging field.

19.10 miRNAs AND DRUG RESPONSE

Aberrant miRNA expression is a common feature of many human diseases. In addition, miRNAs are involved in drug response. For example, the changes in miRNA expression in cancer cells are associated with resistance to anticancer drugs [89], and the underlying mechanism involves the fact that drug transporters (e.g., MDR1/P-glycoprotein, breast cancer resistance protein (BCRP/ABCG2), and multidrug resistance-associated protein 1 (MRP1)) are regulated by miRNAs [90–92]. Furthermore, miRNA expression is affected by chemicals, drugs, and steroid hormones, although these changes are likely smaller than those in mRNA expression. The miRNA manipulation therapy described above would result in larger changes in miRNA expression. miRNAs could be utilized to selectively modulate drug-metabolizing enzymes, drug transporters, and nuclear receptors toward optimal drug responses in pharmacotherapy.

19.11 CONCLUSIONS

There is considerable interindividual variability in the expression and activity of P450s, which consequently results in variability in drug response and toxicity. To understand the reasons behind this phenomenon, considerable effort has been put forth to uncover the mechanisms by which

P450s are transcriptionally and epigenetically regulated. Currently, miRNAs can be recognized as critical regulators of P450s to control pharmacological and toxicological outcomes. Robust studies of miRNA-related polymorphisms will provide a better understanding of interindividual variability in drug response and xenobiotic toxicity. In-depth understanding will be useful for individualized medicine. Future *in vivo* studies on the effects of changes in miRNA expression or function on drug metabolism and pharmacokinetics will offer valuable information for pharmacotherapy.

GLOSSARY

Noncoding RNA a biologically functional RNA actively transcribed from a gene that contains no open reading frame.

miRISC an RNA-induced silencing complex that contains mature miRNA and guides its inhibitory function to the target mRNA.

Argonaute 2 a protein that binds to small RNAs and is a component of RISC. In mammals, there are four Ago proteins, Ago1–4, which function in translational repression; only Ago2 has slicer activity.

Dicer an endoribonuclease that cleaves pre-miRNAs into 20–25-nucleotide, double-stranded miRNAs.

Single-nucleotide polymorphism a single-nucleotide base pair that varies between individuals. SNPs occur approximately once every 100–300 bp in the human genome and constitute an immense source of heterogeneity.

Ortholog a corresponding gene in different species that has descended from the same ancestral gene.

Paralog a gene that is derived from the same ancestral gene but reside in different locations in the same genome.

Cytochrome P450s the human cytochrome P450 superfamily, which comprises 57 genes. P450 sequences that display greater than 40% amino acid identity are classified in the same family; sequences that have greater than 55% identity are classified in the same subfamily. Members of CYP1–3 subfamilies are involved in the metabolism of xenobiotics, including drugs.

Xenobiotic a foreign chemical, which can be manufactured or natural, such as a drug, an industrial chemical, a pollutant, a pyrolysis product in cooked food, an alkaloid, or a secondary plant metabolite.

ABBREVIATIONS

AGTR	angiotensin receptor 1
Ago2	argonaute 2
AhR	aryl hydrocarbon receptor
ARNT	aryl hydrocarbon receptor nuclear translocator
AsO	antisense oligonucleotide
CYP	cytochrome P450
DHFR	dihydrofolate reductase
EET	epoxyeicosatrienoic acid
ER	estrogen receptor
FGF20	fibroblast growth factor 20
HIF-1α	hypoxia inducible factor 1α
MDR1	multidrug resistance 1
miRNA	microRNA
MRE	microRNA recognition element
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
PAHs	polycyclic aromatic hydrocarbons
PANK	pantothenate kinase
PCR-RFLP	polymerase chain reaction–restriction fragment-length polymorphism
PGC1β	peroxisome proliferator-activated receptor γ coactivator 1β
PARα	peroxisome proliferator-activated receptor α
PXR	pregnane X receptor
Pre-miRNA	precursor miRNA
Pri-miRNA	primary miRNA
RISC	RNA-induced silencing complex
SNP	single-nucleotide polymorphism
SULT	sulfotransferase
TRBP	TAR RNA binding protein
UTR	untranslated region
VDR	vitamin D receptor

REFERENCES

[1] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.

[2] Kloosterman WP, Plasterk RH. The diverse functions of microRNAs in animal development and disease. Dev Cell 2006;11:441–50.

[3] Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet 2010;11:597–610.

[4] Havens MA, Reich AA, Duelli DM, Hastings ML. Biogenesis of mammalian microRNAs by a non canonical processing pathway. Nucleic Acids Res 2012;40:4626–40.

[5] Zhang Z, Qin YW, Brewer G, Jing Q. MicroRNA degradation and turnover: regulating the regulators. Wiley Interdiscip Rev RNA 2012;3:593–600.

[6] Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006;34(Database issue):D140–4.

[7] Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem 2010;79:351–79.

[8] Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 2009;19:92–105.

[9] Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005;120:15–20.

[10] Watanabe Y, Tomita M, Kanai A. Computational methods for microRNA target prediction. Methods Enzymol 2007;427:65–86.

[11] Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG. Lost in translation: an assessment and perspective for computational microRNA target identification. Bioinformatics 2009;25:3049–55.

[12] Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 2005;433:769–73.

[13] Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. Nature 2008;455:64–71.

[14] Zhang SY, Surapureddi S, Coulter S, Ferguson SS, Goldstein JA. Human CYP2C8 is post-transcriptionally regulated by microRNAs 103 and 107 in human liver. Mol Pharmacol 2012;82:529–40.

[15] Finnerty JR, Wang WX, Hébert SS, Wilfred BR, Mao G, Nelson PT. The miR-15/107 group of microRNA genes: evolutionary biology, cellular functions, and roles in human diseases. J Mol Biol 2010;402:491–509.
[16] Bièche I, Narjoz C, Asselah T, Vacher S, Marcellin P, Lidereau, et al. Reverse transcriptase-PCR quantification of mRNA levels from cytochrome (CYP)1, CYP2 and CYP3 families in 22 different human tissues. Pharmacogentet Genomics 2007;17:731–42.
[17] Sumida A, Kinoshita K, Fukuda T, Matsuda H, Yamamoto I, Inaba T, et al. Relationship between mRNA levels quantified by reverse transcription-competitive PCR and metabolic activity of CYP3A4 and CYP2E1 in human liver. Biochem Biophys Res Commun 1999;262:499–503.
[18] Mohri T, Nakajima M, Fukami T, Takamiya M, Aoki Y, Yokoi T. Human CYP2E1 is regulated by miR-378. Biochem Pharmacol 2010;79:1045–52.
[19] Wang Z, Hall SD, Maya JF, Li L, Asghar A, Gorski JC. Diabetes mellitus increases the *in vivo* activity of cytochrome P450 2E1 in humans. Br J Clin Pharmacol 2003;55:77–85.
[20] Crunkhorn S, Dearie F, Mantzoros C, Gami H, da Silva WS, Espinoza D, et al. Peroxisome proliferator activator receptor γ coactivator-1 expression is reduced in obesity: potential pathogenic role of saturated fatty acids and p38 mitogen-activated protein kinase activation. J Biol Chem 2007;282:15439–50.
[21] Liu HY, Yehuda-Shnaidman E, Hong T, Han J, Pi J, Liu Z, et al. Prolonged exposure to insulin suppresses mitochondrial production in primary hepatocytes. J Biol Chem 2009;284:14087–95.
[22] De Waziers I, Garlatti M, Bouguet J, Beaune PH, Barouki R. Insulin down-regulates cytochrome P450 2B and 2E expression at the post-transcriptional level in the rat hepatoma cell line. Mol Pharmacol 1995;47:474–9.
[23] Woodcroft KJ, Hafner MS, Novak RF. Insulin signaling in the transcriptional and posttranscriptional regulation of CYP2E1 expression. Hepatology 2002;35:263–73.
[24] Carrer M, Liu N, Grueter CE, Williams AH, Frisard MI, Hulver MW, et al. Control of mitochondrial metabolism and systemic energy homeostasis by microRNAs 378 and 378*. Proc Natl Acad Sci USA 2012;109:15330–5.
[25] Floyd MD, Gervasini G, Masica AL, Mayo G, George Jr AL, Bhat K, et al. Genotype-phenotype associations for common CYP3A4 and CYP3A5 variants in the basal and induced metabolism of midazolam in European- and African-American men and women. Pharmacogenetics 2003;13:596–606.
[26] Takagi S, Nakajima M, Mohri T, Yokoi T. Post-transcriptional regulation of human pregnane X receptor by micro-RNA affects the expression of cytochrome P450 3A4. J Biol Chem 2008;283:9674–80.
[27] Pan Y-Z, Gao W, Yu A-M. MicroRNAs regulate CYP3A4 expression via direct and indirect targeting. Drug Metab Dispos 2009;37:2112–7.
[28] Choi YM, An S, Lee EM, Kim K, Choi SJ, Kim JS, et al. CYP1A1 is a target of miR-892a-mediated post-transcriptional repression. Int J Oncol 2012;41:331–6.
[29] Oda Y, Nakajima M, Mohri T, Takamiya M, Aoki Y, Fukami T, et al. Aryl hydrocarbon receptor nuclear translocator in human liver is regulated by miR-24. Toxicol Appl Pharmacol 2012;260:222–31.
[30] Tsuchiya Y, Nakajima M, Takagi S, Taniya T, Yokoi T. MicroRNA regulates the expression of human cytochrome P450 1B1. Cancer Res 2006;66:9090–8.
[31] Devlin AH, Thompson P, Robson T, McKeown SR. Cytochrome P450 1B1 mRNA untranslated regions interact to inhibit protein translation. Mol Carcinog 2010;49:190–9.
[32] Tsuchiya Y, Nakajima M, Kyo S, Kanaya T, Inoue M, Yokoi T. Human CYP1B1 is regulated by estradiol via estrogen receptor. Cancer Res 2004;64:3119–25.
[33] Adams BD, Furneaux H, White BA. The micro-ribonucleic acid (miRNA) miR-206 targets the human estrogen receptor-α (ERα) and represses ERalpha messenger RNA and protein expression in breast cancer cell lines. Mol Endocrinol 2007;21:1132–47.
[34] Zhao JJ, Lin J, Yang H, Kong W, He L, Ma X, et al. MicroRNA-221/222 negatively regulates estrogen receptor α and is associated with tamoxifen resistance in breast cancer. J Biol Chem 2008;283:31079–86.
[35] Xiong J, Yu D, Wei N, Fu H, Cai T, Huang Y, et al. An estrogen receptor α suppressor, microRNA-22, is downregulated in estrogen receptor α-positive human breast cancer cell lines and clinical samples. FEBS J 2010;277:1684–94.
[36] Jalas JR, Ding X, Murphy SE. Comparative metabolism of the tobacco-specific nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol by rat cytochrome P450 2A3 and human cytochrome P450 2A13. Drug Metab Dispos 2003;31:1199–202.
[37] Kalscheuer S, Zhang X, Zeng Y, Upadhyaya P. Differential expression of microRNAs in early-stage neoplastic transformation in the lungs of F344 rats chronically treated with the tobacco carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. Carcinogenesis 2008;29:2394–9.
[38] Corney DC, Flesken-Nikitin A, Godwin AK, Wang W, Nikitin AY. MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. Cancer Res 2007;67:8433–8.
[39] Enayetallah AE, French RA, Thibodeau MS, Grant DF. Distribution of soluble epoxide hydrolase and of cytochrome P450 2C8, 2C9, and 2J2 in human tissues. J Histochem Cytochem 2004;52:447–54.
[40] Chen F, Chen C, Yang S, Gong W, Wang Y, Cianflone K, et al. Let-7b inhibits human cancer phenotype by targeting cytochrome P450 epoxygenase 2J2. PLoS One 2012;7:e39197.
[41] Jiang JG, Chen CL, Card JW, Yang S, Chen JX, Fu XN, et al. Cytochrome P450 2J2 promotes the neoplastic phenotype of carcinoma cells and is up-regulated in human tumors. Cancer Res 2005;65:4707–15.
[42] Xu S, Linher-Melville K, Yang BB, Wu D, Li J. Micro-RNA378 (miR-378) regulates ovarian estradiol production by targeting aromatase. Endocrinology 2011;152:3941–51.
[43] Panda H, Chuang TD, Luo X, Chegini N. Endometrial miR-181a and miR-98 expression is altered during transition from normal into cancerous state and target PGR, PGRMC1, CYP19A1, DDX3X, and TIMP3. J Clin Endocrinol Metab 2012;97:E1316–26.
[44] Shibahara Y, Miki Y, Onodera Y, Hata S, Chan MS, Yiu CC, et al. Aromatase inhibitor treatment of breast cancer cells increases the expression of let-7f, a microRNA targeting CYP19A1. J Pathol 2012;227:357–66.
[45] Carlberg C, Polly P. Gene regulation by vitamin D_3. Crit Rev Eukaryot Gene Expr 1998;8:19–42.
[46] Deeb KK, Trump DL, Johnson CS. Vitamin D signaling pathways in cancer: potential for anticancer therapeutics. Nat Rev Cancer 2007;7:684–700.
[47] Friedrich M, Axt-Fliedner R, Villena-Heinsen C, Tilgen W, Schmidt W, Reichrath J. Analysis of vitamin D-receptor (VDR) and retinoid X-receptor α in breast cancer. Histochem J 2002;34:35–40.

[48] Khadzkou K, Buchwald P, Westin G, Dralle H, Akerström G, Hellman P. 25-Hydroxyvitamin D_3 1α-hydroxylase and vitamin D receptor expression in papillary thyroid carcinoma. J Histochem Cytochem 2006;54:355–61.

[49] Komagata S, Nakajima M, Takagi S, Mohri T, Taniya T, Yokoi T. Human CYP24 catalyzing the inactivation of calcitriol is post-transcriptionally regulated by miR-125b. Mol Pharmacol 2009;76:702–9.

[50] Mohri T, Nakajima M, Takagi S, Komagata S, Yokoi T. MicroRNA regulates human vitamin D receptor. Int J Cancer 2009;125:1328–33.

[51] Scott GK, Goga A, Bhaumik D, Berger CE, Sullivan CS, Benz CC. Coordinate suppression of ERBB2 and ERBB3 by enforced expression of microRNA miR-125a or miR-125b. J Biol Chem 2007;282:1479–86.

[52] Li W, He X, Li J, Tu K, Wei L, Wu J, et al. Diagnostic and prognostic implications of microRNAs in human hepatocellular carcinoma. Int J Cancer 2008;123:1616–22.

[53] Visone R, Pallante P, Vecchione A, Cirombella R, Ferracin M, Ferraro A, et al. Specific microRNAs are downregulated in human thyroid anaplastic carcinomas. Oncogene 2007;26:7590–5.

[54] Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci USA 2006;103:2257–61.

[55] Moffat ID, Boutros PC, Celius T, Lindén J, Pohjanvirta R, Okey AB. MicroRNAs in adult rodent liver are refractory to dioxin treatment. Toxicol Sci 2007;99:470–87.

[56] Shah YM, Morimura K, Yang Q, Tanabe T, Takagi M, Gonzalez FJ. Peroxisome proliferator-activated receptor α regulates a microRNA-mediated signaling cascade responsible for hepatocellular proliferation. Mol Cell Biol 2007;27:4238–47.

[57] Fukushima T, Hamada Y, Yamada H, Horii I. Changes of micro-RNA expression in rat liver treated by acetaminophen or carbon tetrachloride, regulation role of micro-RNA for RNA expression. J Toxicol Sci 2007;32:401–9.

[58] Pogribny IP, Tryndyak VP, Boyko A, Rodriguez-Juarez R, Beland FA, Kovalchuk O. Induction of microRNAome deregulation in rat liver by long-term tamoxifen exposure. Mutat Res 2007;619:30–7.

[59] Murakami T, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, et al. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. Oncogene 2006;25:2537–45.

[60] Zhang Y, Wu L, Wang Y, Zhang M, Li L, Zhu D, et al. Protective role of estrogen-induced miRNA-29 expression in carbon tetrachloride-induced mouse liver injury. J Biol Chem 2012;287:14851–62.

[61] Roderburg C, Urban GW, Bettermann K, Vucur M, Zimmermann H, Schmidt S, et al. Micro-RNA profiling reveals a role for miR-29 in human and murine liver fibrosis. Hepatology 2011;53:209–18.

[62] Saunders MA, Liang H, Li WH. Human polymorphism at microRNAs and microRNA target sites. Proc Natl Acad Sci USA 2007;104:3300–5.

[63] Sethupathy P, Collins FS. MicroRNA target site polymorphisms and human disease. Trends Genet 2008;24:489–97.

[64] Chin LJ, Ratner E, Leng S, Zhai R, Nallur S, Babar I, et al. A SNP in a *let-7* microRNA complementary site in the *KRAS* 3′ untranslated region increases non-small cell lung cancer risk. Cancer Res 2008;68:8535–40.

[65] Abelson JF, Kwan KY, O'Roak BJ, Baek DY, Stillman AA, Morgan TM, et al. Sequence variants in SLITRK1 are associated with Tourette's syndrome. Science 2005;310:317–20.

[66] Sethupathy P, Borel C, Gagnebin M, Grant GR, Deutsch S, Elton TS, et al. Human microRNA-155 on chromosome 21 differentially interacts with its polymorphic target in the AGTR1 3′ untranslated region: a mechanism for functional single-nucleotide polymorphisms related to phenotypes. Am J Hum Genet 2007;81:405–13.

[67] Wang G, van der Walt JM, Mayhew G, Li YJ, Züchner S, Scott WK, et al. Variation in the miRNA-433 binding site of FGF20 confers risk for Parkinson disease by overexpression of α-synuclein. Am J Hum Genet 2008;82:283–9.

[68] Kapeller J, Houghton LA, Mönnikes H, Walstab J, Möller D, Bönisch H, et al. First evidence for an association of a functional variant in the microRNA-510 target site of the serotonin receptor-type 3E gene with diarrhea predominant irritable bowel syndrome. Hum Mol Genet 2008;17:2967–77.

[69] Mishra PJ, Humeniuk R, Mishra PJ, Longo-Sorbello GS, Banerjee D, Bertino JR. A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. Proc Natl Acad Sci USA 2007;104:13513–8.

[70] Hayashi H, Tazoe Y, Horino M, Fujimaki-Katoh C, Tsuboi S, Matsuyama T, et al. An artifact derived from a pseudogene led to the discovery of microRNA binding site polymorphism in the 3′-untranslated region of the human dihydrofolate reductase gene. Drug Metab Pharmacokinet 2012;27:263–7.

[71] Goto Y, Yue L, Yokoi A, Nishimura R, Uehara T, Koizumi S, et al. A novel single-nucleotide polymorphism in the 3′-untranslated region of the human dihydrofolate reductase gene with enhanced expression. Clin Cancer Res 2001;7:1952–6.

[72] Chen MJ, Shimada T, Moulton AD, Harrison M, Nienhuis AW. Intronless human dihydrofolate reductase genes are derived from processed RNA molecules. Proc Natl Acad Sci USA 1982;79:7435–9.

[73] Yu X, Dhakal IB, Beggs M, Edavana VK, Williams S, Zhang X, et al. Functional genetic variants in the 3′-untranslated region of sulfotransferase isoform 1A1 (SULT1A1) and their effect on enzymatic activity. Toxicol Sci 2010;118:391–403.

[74] Nakajima M, Kwon JT, Tanaka N, Zenta T, Yamamoto Y, Yamamoto H, et al. Relationship between interindividual differences in nicotine metabolism and CYP2A6 genetic polymorphism in humans. Clin Pharmacol Ther 2001;69:72–8.

[75] Wang J, Pitarque M, Ingelman-Sundberg M. 3′-UTR polymorphism in the human CYP2A6 gene affects mRNA stability and enzyme expression. Biochem Biophys Res Commun 2006;340:491–7.

[76] Liao LH, Zhang H, Lai MP, Chen SL, Wu M, Shen N. Single-nucleotide polymorphisms and haplotype of CYP2E1 gene associated with systemic lupus erythematosus in Chinese population. Arthritis Res Ther 2011;13:R11. http://dx.doi.org/10.1186/ar3232.

[77] Ramamoorthy A, Skaar TC. In silico identification of microRNAs predicted to regulate the drug metabolizing cytochrome P450 genes. Drug Metab Lett 2011;5:126–31.

[78] Wei R, Yang F, Urban TJ, Li L, Chalasani N, Flockhart DA, et al. Impact of the interaction between 3′-UTR SNPs and microRNA on the expression of human xenobiotic metabolism enzyme and transporter genes. Front Genet 2012;3:248. http://dx.doi.org/10.3389/fgene.2012.00248.

[79] Iwai N, Naraba H. Polymorphisms in human pre-miRNAs. Biochem Biophys Res Commun 2005;331:1439–44.

[80] Duan R, Pak C, Jin P. Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. Hum Mol Genet 2007;16:1124–31.

[81] Gong J, Tong Y, Zhang HM, Wang K, Hu T, Shan G, et al. Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. Hum Mutat 2012;33:254–63.

[82] Hu Z, Liang J, Wang Z, Tian T, Zhou X, Chen J, et al. Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. Hum Mutat 2009;30:79–84.

[83] Arisawa T, Tahara T, Shibata T, Nagasaka M, Nakamura M, Kamiya Y, et al. A polymorphism of microRNA 27a genome region is associated with the development of gastric mucosal atrophy in Japanese male subjects. Dig Dis Sci 2007;52:1691–7.

[84] Jazdzewski K, Murray EL, Franssila K, Jarzab B, Schoenberg DR, de la Chapelle A. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. Proc Natl Acad Sci USA 2008;105:7269–74.

[85] Boni V, Zarate R, Villa JC, Bandrés E, Gomez MA, Maiello E, et al. Role of primary miRNA polymorphic variants in metastatic colon cancer patients treated with 5-fluorouracil and irinotecan. Pharmacogenomics J 2011;11:429–36.

[86] Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857–66.

[87] Seto AG. The road toward microRNA therapeutics. Int J Biochem Cell Biol 2010;42:1298–305.

[88] Wahid F, Shehzad A, Khan T, Kim YY. MicroRNAs: synthesis, mechanism, function, and recent clinical trials. Biochim Biophys Acta 2010;1803:1231–43.

[89] Sarkar FH, Li Y, Wang Z, Kong D, Ali S. Implication of microRNAs in drug resistance for designing novel cancer therapy. Drug Resist Updat 2010;13:57–66.

[90] Zhu H, Wu H, Liu X, Evans BR, Medina DJ, Liu CG, et al. Role of MicroRNA miR-27a and miR-451 in the regulation of MDR1/P-glycoprotein expression in human cancer cells. Biochem Pharmacol 2008;76:582–8.

[91] Pan YZ, Morris ME, Yu AM. MicroRNA-328 negatively regulates the expression of breast cancer resistance protein (BCRP/ABCG2) in human cancer cells. Mol Pharmacol 2009;75:1374–9.

[92] Liang Z, Wu H, Xia J, Li Y, Zhang Y, Huang K, et al. Involvement of miR-326 in chemotherapy resistance of breast cancer through modulating expression of multidrug resistance-associated protein 1. Biochem Pharmacol 2010;79:817–24.

[93] Nakajima M. MicroRNAs regulating metabolism of drugs, xenobiotics and endobiotics. Showa Univ J Pharm Sci 2013;4:1–11.

This page is intentionally left blank

Part VII

Application in Therapeutics

This page is intentionally left blank

Chapter 20

Adverse Drug Reactions

Lauren Walker, Vincent Yip, and Munir Pirmohamed
The Wolfson Centre for Personalized Medicine, Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool, UK

Chapter Outline

20.1 Overview 405
20.1.1 Classification 405
20.1.2 The Burden of ADRs 407
20.1.3 Predisposition to Adverse Drug Reactions 407
20.2 Phase I Metabolism 408
20.2.1 P450 Phenotypes 408
20.3 Phase II Metabolism 410
20.3.1 Thiopurines 410
20.3.2 Cisplatin 411
20.3.3 Isoniazid 411
20.3.4 Sulfamethoxazole 412
20.3.5 Irinotecan 412
20.3.6 Combined PKPD 413
20.4 Type B Adverse Drug Reactions 414
20.4.1 Malignant Hyperthermia and Anesthesia 414
20.4.2 Statins and Skeletal Muscle Toxicity 415
20.4.3 Drug-Induced Skin Injury 417
20.4.4 Drug-Induced Liver Injury 421
20.5 Conclusion 424
Abbreviations and Acronyms 425
References 426

20.1 OVERVIEW

Interindividual variability in drug response represents a major problem for both drug development and clinical practice. Adverse drug reactions (ADRs) are defined as "an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product; adverse effects usually predict hazard from future administration and warrant prevention, or specific treatment, or alteration of the dosage regimen, or withdrawal of the product" [1].

20.1.1 Classification

ADRs were originally classified by Rawlins and Thompson into two basic types: A and B (Table 20.1) [2]. Type A (or augmented) reactions are predictable, usually dose-dependent, and based on an exaggerated response to the known primary and/or secondary pharmacological action(s) of the drug. This is illustrated by bleeding episodes with warfarin therapy or hyperkalemia with potassium-sparing diuretics. Type B (or bizarre) reactions are unpredictable aberrant effects that are not expected from the known pharmacological action(s) of the drug when given in the usual therapeutic doses. This includes anaphylaxis or cutaneous eruptions with penicillins. However, as our knowledge of adverse drug reactions increases, some of these type B reactions will become predictable. Over time, the type A and B classification of ADRs was expanded to include types C, D, and E reactions (Table 20.2) [3].

More recently, a new classification system for ADRs was developed based on dose relatedness, timing, and patient susceptibility (DoTS) (Table 20.3) [4]. This system incorporates not only the properties of the drug (pharmacology and dose), but also the characteristics of the reaction (time course and severity) and the properties of the individual patient (genetic, pathological, or other biological differences that confer susceptibility).

The Rawlins and Thompson classification suggested that all type B ADRs are dose-independent, but this is not the case, as illustrated by type IV hypersensitivity reactions and desensitization using increasing dosages of antigen in cephalosporin allergies [5,6]. According to the DoTS classification, dose relatedness should be divided into reactions that occur at supratherapeutic doses (toxic effects), reactions that occur at standard therapeutic doses (collateral effects), and reactions that occur at subtherapeutic doses in susceptible patients (hypersusceptibility effects). For the latter, in particular, genetic

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00020-7
© 2014 Elsevier Inc. All rights reserved.

TABLE 20.1 Clinical Features and Classification of Type A and B Drug Reactions

Features	Type A	Type B
Pharmacology	Augmented	Bizarre
Predictable	Yes	No
Dose-dependent	Yes	No
Morbidity	High	Low
Mortality	Low	High

TABLE 20.2 Expanded Adverse Drug Reactions Classification

Adverse Drug Reactions	Definition	Examples
Type C	Continuing	Osteonecrosis of the jaw with bisphosphonates
Type D	Delayed	Leucopenia 6 weeks after lomustine dose
Type E	End-of-use	Insomnia, anxiety, and perceptual disturbances following withdrawal of benzodiazepines

TABLE 20.3 DoTS Classification for ADRs

	Variable	Classification
Do	Dose relatedness	Supratherapeutic Standard therapeutic Subtherapeutic
T	Timing	Time-independent Time-dependent
S	Susceptibility	Age Gender Genetics Pathology Physiology Environmental

factors may be important. Timing of ADRs is divided into time-independent and time-dependent. Time-independent reactions can happen at any time during the treatment course, but typically occur when the concentration of the drug at the site of action changes (e.g., renal failure) or when the pharmacological response to the drug is altered without a change in concentration (e.g., digoxin and hypokalemia). Time-dependent reactions are further subdivided into six subtypes: rapid, first dose, early, intermediate, late, and delayed (Table 20.4) [4].

TABLE 20.4 Subtype Classification of Time-Dependent Subtypes in DoTS Classification of ADRs

Subtype	Features	Examples
Rapid	Occurs when drug is administered too rapidly	Red-man syndrome with vancomycin
First dose	Occurs after first dose of treatment course	Hypotension after first dose of angiotensin-converting enzyme inhibitor (ACEi)
Early	Occurs early then abates with further treatment	Nitrate-induced headache
Intermediate	Occurs after some delay	Neutropenia with carbimazole
Late	Increased occurrence with recurrent or continued exposure	Adverse effects of corticosteroids
Delayed	Observed some time after exposure even if drug withdrawn before reaction occurs	Teratogenesis associated with thalidomide

Neither the Rawlins and Thompson nor the DoTS classification system is perfect, but both serve as important guidance for pharmacovigilance in clinical practice and drug regulation while providing important insights for future diagnosis, management, and prevention of ADRs.

20.1.2 The Burden of ADRs

ADRs are a major cause of morbidity, mortality, and increased healthcare costs. A study of almost 19,000 admissions at two National Health Service (NHS) hospitals in the United Kingdom found that 6.5% were related to an ADR [7]. Furthermore, of 3695 inpatient episodes occurring across 12 wards over a 6-month period, ADRs were seen in almost 15% and caused an increase in length of stay by an average of 0.25 days per patient admission [8]. Clearly, this places a significant financial burden on any healthcare system; in the NHS, a conservative figure of the financial burden of ADRs in adults is in the region of £637 million per annum [8]. In the United States, a meta-analysis of 39 prospective studies showed that 6.7% of hospital inpatients suffer serious ADRs, resulting in a mortality estimate of 100,000 U.S. deaths per year [9]. Similar results from a Swedish study showed that ADRs were the seventh most common cause of fatality [10].

A meta-analysis of 102 studies in children identified incidence rates for ADRs causing hospital admission to be 0.4–10.3% of all children (pooled estimate of 2.9% (2.6%, 3.1%)) and 0.6–16.8% of all children exposed to a drug during a hospital stay [11]. In accordance with these data, a prospective study of 8345 patient episodes in a large tertiary children's hospital estimated an incidence rate of 2.9% (95% CI 2.5, 3.3). Most of those serious enough to require hospital admission resulted from hospital-based prescribing, of which over a fifth may have been avoidable [12].

20.1.3 Predisposition to Adverse Drug Reactions

Predisposition, and thereby susceptibility, to ADRs can be due to many factors, including age, gender, genetic variation, pathology, physiological variation, and environmental factors (Table 20.5). It is therefore important to evaluate ADRs on a case-by-case basis and evaluate the role of both environmental and genetic factors. It is difficult to estimate the quantitative contribution of genetic factors to ADRs, which is likely to vary among drugs.

20.1.3.1 Twin Studies

Twin studies are a valuable tool to assess the relative contribution of genetic and environmental factors to variation in complex phenotypes, including drug response. Twin studies compare the similarity of a trait between monozygotic (MZ) twins who share 100% genetic identity and that of dizygotic (DZ) twins who on average share only 50% of the genome. It can be assumed that twins are raised in identical environments; therefore, any greater similarity of a trait shared between MZ twins compared to DZ twins can be attributed to genetics.

One of the first twin studies to identify a pharmacogenetic phenotype for ADR involved the antituberculosis (TB) drug isoniazid. In a study comparing metabolism between five pairs of MZ and DZ twins, investigators showed that the metabolism of isoniazid was far less variable in the former than in the latter [13]. This was later confirmed to be due, at least in part, to two distinct acetylation phenotypes (slow and rapid) that are determined by polymorphisms in the

TABLE 20.5 Sources of Altered Susceptibility to ADRs

Susceptibility Variable	Examples
Age	Increased susceptibility to hepatotoxicity with flucloxacillin[a]
Gender	Increased frequency of cutaneous adverse reactions among females[b]
Genetics	Azathioprine-induced myelosuppression in patients with significantly reduced thiopurine methyltransferase (TPMT) activity[c]
Pathology	Increased risk of lithium toxicity in patients with renal failure[d]
Physiology	Increased antiepileptic clearance among pregnant women with epilepsy associated with increased seizures[e]
Environmental factors	Consumption of grapefruit juice can affect bioavailability of numerous medications metabolized by CYP3A4 or transported by OATP[f]

[a] *Olsson R, Wiholm BE, Sand C, et al. Liver damage from flucloxacillin, cloxacillin and dicloxacillin. J Hepatol 1992;15:154–61.* [b] *Pistone G, Gurreri R, Alaimo R, et al. Gender differences in adverse drug reactions in dermatological patients in west Sicily: an epidemiological study [Epub ahead of print]. J Dermatol Treat 2013. doi: http://dx.doi.org/10.3109/09546634.2013.814755.* [c] *See Ref. [54].* [d] *McKnight RF, Adida M, Budge, K, et al. Lithium toxicity profile: a systematic review and meta-analysis. Lancet 2012;379:721–8.* [e] *Reisinger TL, Newman M, Loring DW, et al. Antiepileptic drug clearance and seizure frequency during pregnancy in women with epilepsy. Epilepsy Behav 2013;29: 13–8.* [f] *Pirmohamed M. Drug–grapefruit juice interactions. BMJ 2013;346:f1.*

N-acetyltransferase gene [14]. We now know that individuals who are "slow acetylators" are at greater risk of toxicity with compounds such as sulphonamides and aromatic amine carcinogens, and with isoniazid, which leads to hepatic toxicity and peripheral neuropathy [15–17]. Twin studies with other drugs, including antipyrine, phenytoin, halothane, and phenylbutazone have all shown that genetic factors are the main determinants of interindividual variation in metabolism [18]. However, in the majority of cases, it is likely that variability is multigenetic in origin and thus poses challenges in developing predictive tests [19].

20.2 PHASE I METABOLISM

20.2.1 P450 Phenotypes

It is likely that CYP450 polymorphisms have the greatest impact on type A ADRs where a single gene product predominantly determines drug clearance and consequently toxicological response. Polymorphisms affecting most of the genes encoding cytochrome P450 enzymes have been described. Variability in enzyme expression, or alteration in substrate specificity, may lead to interindividual variability in drug metabolism, which in turn affects drug plasma levels and thus susceptibility to adverse events. Poor enzyme activity may result in toxic drug accumulation, just as ultrarapid metabolism may cause rapid accumulation of active and, in some cases, more reactive metabolites. The genotype–phenotype correlation can be translated into four major phenotypes (Table 20.6).

20.2.1.1 CYP2D6

By far, the most intensively studied CYP450 gene is CYP2D6, which is responsible for the metabolism of 25% of all drugs [22]. To date, more than 70 variant alleles and over 130 genetic variants at the CYP2D6 locus have been described [23]. As a result of one or more point mutations, gene deletions, or multiplications, the resultant mutations can either have no effect on enzyme activity or can code for enzymes with decreased/absent activity (e.g., _CYP2D6*10_) or increased activity (e.g., _CYP2D6*2xN_) [20]. Absent or decreased CYP2D6 activity can lead to ADRs through either reduced first-pass metabolism, accumulation of the drug following reduced metabolism, or rerouting of metabolism. Similarly, increased activity may lead to ADRs through increased formation of active metabolites.

β-Blockers

Metoprolol undergoes extensive hepatic first-pass metabolism [24] with CYP2D6 the major enzyme involved [25,26]. Carriers of the poor metabolizer (PM) phenotype display significant increases in the bioavailability and half-life of the drug, with consequent increases in both peak and steady-state concentrations: metoprolol elimination half-life increases from 2.8 h in extensive metabolizers (EM) to approximately 7.5 h in poor metabolizers [26]. Furthermore, median dose-corrected plasma concentrations of metoprolol have been shown to be 6.2- and 3.9-fold higher in individuals carrying the PM or intermediate metabolizer (IM) phenotypes, respectively, than in EMs [27]. An increased risk of toxicity after metoprolol administration has been demonstrated, with PMs reported to have a 5-fold higher risk of ADRs [28]. Some prospective studies in hypertension [29] and heart failure [30], however, have failed to show an association between CYP2D6 metabolizer status and the occurrence of ADRs associated with metoprolol. However, these studies contained only small numbers of PMs (4 and 8, respectively) and must therefore be interpreted with caution.

CYP2D6 is also involved in the metabolism of other β-blockers, and the higher risk of toxicity in PMs has been demonstrated after administration of timolol (more pronounced bradycardia) [31] and carvedilol (reduction of

TABLE 20.6 Genotype–Phenotype Correlation of Cytochrome P450 Drug-Metabolizing Enzymes

Metabolizer Phenotype	Defect	Frequency	Consequence	Plasma Drug Level	Mechanism	Individualized Therapy
Poor (PM)	Carrying two defective alleles	5–10% [20,21][a]	Completely lack enzyme activity	Toxic	Creates unstable enzymes or amino acid changes leading to reduced substrate binding	Decrease dose to decrease toxicity
Intermediate (IM)	Heterozygous for a defective allele	[10–15]	Decreased enzyme activity	High		Standard dosing
Extensive (EM)	Carrying two functional alleles	75–85%	Proportionally increased amount of enzyme expressed in relation to number of gene copies. Promoter region and regulatory elements controlling gene transcription are duplicated	Normal	Increased enzyme expression with increased production of active metabolite	Standard dosing
Ultrarapid (URM)	Carrying >2 active gene copies (e.g., CYP2D6) Gain-of-function polymorphism (e.g., *CYP2C19*17*)	1–10% [20,21]		Low		Increase dose to increase efficacy

[a] *See references at the end of the chapter publications cited.*
Source: Adapted from Pharmacokinetics: Concepts and Applications in Drug Discovery and Development. John Wiley and Sons (forthcoming).

systolic blood pressure) [32]. Although known perturbations in PK parameters caused by PM status affect plasma concentrations of metoprolol and carvedilol, and hence can reasonably be assumed to result in greater adverse responses, larger, long-term prospective studies are required. These should include greater numbers of PMs in order to fully characterize whether this association is clinically significant.

Codeine

In contrast, ultrarapid metabolizers are at higher risk of codeine ADRs as a result of greater O-demethylation to codeine's active metabolite, morphine. The UM phenotype is associated with 50% greater plasma morphine concentrations than are EM phenotypes [33]. In February 2013, the U.S. food and drug administration (FDA) introduced a new Boxed Warning and Contraindication to the drug label of codeine-containing products [34]. This relates to the reporting of three fatalities and one life-threatening case of respiratory depression in children with the UM phenotype who were given codeine following tonsillectomy or adenoidectomy [35,36]. This triggered a European review of the safety of codeine-containing medicines in children. As a result, the MHRA issued a safety update in July 2013 stating that codeine is now contraindicated in children under 18 undergoing tonsillectomy or adenoidectomy for the treatment of obstructive sleep apnea. Codeine use is now restricted to children over the age of 12 for short-lived moderate pain unrelieved by other painkillers [37].

The latest developments in the restriction of codeine use followed a case report in 2006 that detailed the death of a neonate following the use of codeine for episiotomy pain following childbirth by the mother, who had the UM phenotype. After the mother's use of codeine (30 mg) and acetaminophen (500 mg) two tablets every 12 h (reduced by 50% on day two because of constipation), the infant displayed initial lethargy and poor breastfeeding, but by day 12 his skin was grey and he died on day 13 as a result of opioid overdose [38]. His postmortem blood morphine level was 70 ng/mL. The typical serum concentration in neonates breastfed by mothers taking regular codeine is usually 0–2.2 ng/mL [39]. This has led to safety warnings about the use of codeine in breast-feeding mothers by many regulatory authorities; however, regulatory bodies have not recommended genetic testing for codeine prescribing, although they do advise close monitoring, particularly in high-risk populations.

20.2.1.2 CYP2C9

Phenytoin

Phenytoin is the most widely prescribed antiepileptic drug in the U.S., accounting for 52% of prescriptions [40]. It has a narrow therapeutic index and shows wide interindividual variation. It is therefore difficult to administer effectively, requiring therapeutic drug monitoring. Phenytoin is metabolized to a mixture of (*R*) and (*S*) stereoisomers of the inactive compound 5-(4′-hydroxyphenyl)-5-phenylhydantoin (p-HPPH) by CYP2C9 and CYP2C19. CYP2C9 accounts for up to 90% of the metabolism of PHT [41]. Numerous previous studies demonstrated that the presence of *CYP2C9*2* and **3* alleles result in reduced *in vivo* phenytoin metabolism and reduced elimination, with an increase in drug concentrations necessitating the use of lower doses in patients [42–45]. In a recent candidate gene study [46] involving 182 Caucasians with epilepsy, 49 reported ADRs that were significantly associated with the presence of *CYP2C9*2* and **3* alleles. The association was replicated in a further 817 epilepsy patients. However, phenotypic data was collected retrospectively and therefore a number of confounding variables, including drug–drug interactions and compliance, may have led to misclassification. In another study of 281 epilepsy patients, the *CYP2C9*3* allele showed a significant association with maximum phenytoin dose but did not correlate with ADRs [47]. Interestingly, in a small study of 19 epileptic patients with the *CYP2C9*2* and **3* variants treated with phenytoin for over one year, there was significantly reduced cerebellar white matter volume in the variant group compared with demographically matched wild-type controls [48]. The relevance of this is unclear. Large prospective studies with adequate adjustment for confounding are required to assess the contribution of CYP2C9 polymorphisms and the risk of ADRs with phenytoin.

20.2.1.3 Other CYP450 Enzymes

Tardive Dyskinesia and Antipsychotics

Polymorphisms affecting response to antipsychotic medications are an example of where it has been difficult to replicate initial positive findings. There is wide variability in terms of response and tolerability with both typical and atypical antipsychotic drugs. Approximately 40% of antipsychotics are major substrates of CYP2D6, 23% are substrates of CYP3A4, and 18% are substrates of CYP1A2 [49]. This provides a plausible biological rationale for potentially explaining the degree of variability in response. However, a recent meta-analysis failed to show a significant association between tardive dyskinesia and CYP2D6 or CYP1A2 overall [50]. Still, restricting the analysis to prospective studies revealed a significant association between tardive dyskinesia and CYP2D6 in those who carried either two copies (OR = 2.08, 95% CI 1.21–3.57) or one copy (OR = 1.83, 95% CI 1.09–3.08) of the variant alleles when compared with the wild-type. In a large study of 100 patients taking psychotropic medications, Chou and colleagues [51] identified an increased risk of ADRs in PMs associated with prolonged hospital stay and greater treatment cost. Furthermore, a study of 85 patients with schizophrenia showed that patients with the C/C genotype for *CYP1A2*1F* are at significantly increased risk of tardive dyskinesia induced by typical antipsychotics, compared with those who are heterozygous or homozygous for the *CYP1A2*A* allele ($p = 0.0007$) [52]. However, Schulze and colleagues [53] failed to replicate this association in 119 German patients with schizophrenia. Patient heterogeneity and poor phenotyping may explain the disparity to a degree; this is a problem that affects all genetic case-control association studies.

To date, no large randomized control trials (RCTs) have evaluated whether genotyping can successfully reduce the frequency of ADRs associated with antipsychotic medication and the consequent treatment interruption or withdrawal—which has a large impact on outcome. This information is essential for cost-effectiveness analysis and implementation in clinical practice, but of course large RCTs are expensive, particularly given how infrequent some polymorphisms are.

20.3 PHASE II METABOLISM

20.3.1 Thiopurines

Thiopurine S-Methyltransferase (TPMT) catalyzes the S-methylation of thiopurine drugs, including azathioprine (AZA) and 6-mercaptopurine (6-MP). Both are prodrugs metabolized via a series of steps to the 6-thioguanines. These are then incorporated into DNA and RNA, causing immunosuppression. AZA and 6-MP are suitable treatments for a range of immune diseases including rheumatoid arthritis, inflammatory bowel disease, and pediatric acute lymphoblastic leukemia (ALL).

20.3.1.1 Adverse Effects of Thiopurines

A serious adverse effect of both AZA and 6-MP is severe bone marrow depression [54,55]. This results from deficiency of the TPMT enzyme and leads to the increased formation of a cytotoxic 6-thioguanosine nucleotide metabolite via an alternative metabolism pathway. In addition to bone marrow toxicity, hepatitis and pancreatitis are among the most common ADRs associated with thiopurine use, leading to dose reduction or discontinuation in 9–28% of patients [56]. ADRs can occur with the thiopurines at both ends of the enzymic-activity spectrum. Those with low TPMT activity effectively receive an overdose with standard dosing, and hence are 30–40% more likely to develop an ADR. Conversely, very high TPMT activity increases the risk of hepatotoxicity resulting from increased production of methylated thiopurine metabolites [57,58].

20.3.1.2 TPMT Phenotype

The TPMT phenotype test quantifies the level of enzyme activity in red blood cells. Phenotyping is considered to be more reliable for the detection of myelosuppression risk [59,60]. However, the retrospective nature of the relevant studies may not have provided a true measure of ADRs. Recent reports are emerging that suggest genotyping to be the most reliable test, particularly in heterozygotes [61].

20.3.1.3 TPMT Genotype

Approximately 86–97% of patients have the *TPMT*1/*1* (wild-type) genotype, conferring normal TPMT enzyme activity. *TPMT*3C* contains A719G on exon 10 [62], *TPMT*3A* is characterized by a double mutant allele (G460A on exon 7 and A719G on exon 10) [62–66], and *TPMT*2* is a G238C transition on exon 5 [63]. Together these polymorphisms account for 80–95% of deficient TPMT activity in Caucasians, African-Americans, and Asians. Approximately 3–14% of patients are heterozygous for the *TPMT* genotype and possess one variant allele [67]. Deficiency is associated with homozygosity for these alleles. The population prevalence for deficiency is low (approximately 1 in 178 to 1 in 3736 patients) [67], but these patients are at greatest risk for severe myelosuppression and associated life-threatening complications, including sepsis. It has therefore been suggested that dose reduction or alternative drug selection may be appropriate in heterozygotes and homozygotes, respectively.

In the largest prospective study to use full-dose azathioprine (2 mg/kg) without dose adjustment in patients with IBD, Ansari and colleagues found that the heterozygous *TPMT* genotype strongly predicted ADRs (79% heterozygous versus 35% wild-type TPMT, $p<0.001$), especially for gastric intolerance and myelosuppression [68].

Currently there are eight genotype tests available and no single gold standard. Systematic review of all methods in a nonspecific population concluded that genotype tests overall have lower sensitivity but greater positive predictive value than the TPMT phenotype test, but a wide range of sensitivity and specificity [69]. The *TPMT* genotype test cannot be considered the gold standard for identification of myelosuppression risk because the genotype is not solely responsible for this risk and the test itself is designed to detect a limited number of mutations. In clinical practice, it is estimated that only approximately 25% of cases of leukopenia in patients with IBD treated with thiopurines are related to the presence of a polymorphism [70].

20.3.1.4 Perspectives of Professional and Regulatory Bodies

The FDA and prescribing information for AZA and 6-MP recommend *TPMT* genotyping or phenotyping prior to initiating therapy to improve identification of at-risk patients [71]. The American College of Gastroenterology treatment guidelines prefer *TPMT* phenotyping over genotyping because phenotyping reports a quantitative level of TPMT enzyme activity [72]. In the United Kingdom, phenotyping is the preferred method for routine TPMT screening. However, the British Society of Gastroenterology guidelines do not require either *TPMT* genotyping or *TPMT* phenotyping prior to initiating thiopurine therapy because the use of AZA has been shown to be relatively safe in patients with IBD, and the occurrence of leucopenia can be handled through regular blood count monitoring [73,74].

20.3.1.5 Cost-Effectiveness

Six economic evaluations conducted in rheumatological disorders and inflammatory bowel disease have concluded that pretreatment TPMT testing for AZA is a cost-effective use of healthcare resources [75–80]. Genotype-directed dosing schedules compared favorably with other dosing schedules, with pretreatment genotyping costing $579 per life-year saved for a 30-year-old and $1364 per life-year saved for a 60-year-old [75]. The studies reporting these results were retrospective modeling studies that synthesized existing clinical and cost data supplemented by expert opinion where necessary. However, a recent cost-effectiveness analysis in children with acute lymphoblastic leukemia (ALL) reported that screening for TPMT mutations using either genotype or enzymatic laboratory tests prior to administration of 6-MP in pediatric ALL patients did not demonstrate any survival benefit and was therefore not cost-effective. Incremental costs per child (95% confidence interval) were $277 and $298 for the genotyping and phenotyping strategies, respectively, compared to weight-based dosing [81].

20.3.2 Cisplatin

TPMT deficiency related genotypes have also been identified as a risk factor for drug-induced ototoxicity (hearing loss) in children treated with cisplatin for a range of neoplasms [82]. Indeed, in this study 26 of 162 children carried one or more *TPMT* variants, and of these 26, 96% [23] suffered from the most severe ototoxicity. This finding was recently replicated in 155 children and led to a revision of the cisplatin drug label by the FDA to include information relating to *TPMT* pharmacogenetics and ototoxicity [83,84].

20.3.3 Isoniazid

Isoniazid is among the first line antituberculosis (TB) drugs, commonly given in combination with rifampicin and pyrazinamide as RIFATER®. Metabolism and excretion of isoniazid is principally by N-acetyltransferases (NAT) 2-mediated acetylation and CYP2E1.

NATs catalyze the biotransformation of aromatic amines and hydrazines by transferring the acetyl group of acetyl coenzyme A to the free amino group of the parent compound. They are found in two forms in humans: NAT1 and NAT2. NAT1 is ubiquitous, whereas NAT2 expression is generally restricted to the liver and intestine. NAT2 is highly polymorphic, with several mutations affecting its activity. The mutations result in three main acetylator phenotypes: (1) slow acetylators resulting from amino acid changes include *NAT-2*5*, *NAT-2*6*, *NAT-2*7*, and *NAT-2*14* [85]; (2) homozygous alleles at all four loci produce the rapid acetylator phenotype, with heterozygotes referred to as intermediate metabolizers [85–88]; (3) NAT2-determined slow metabolizers acetylate isoniazid slower than fast metabolizers, resulting in the accumulation of toxic intermediates in the liver and leading to hepatotoxicity. This has been demonstrated in those with the slow acetylator phenotype [89], although there is some conflicting evidence [90].

Recently, rs1495741, a novel tag SNP in NAT2 was reported to confer high sensitivity for prediction of NAT2 acetylator phenotype in Europeans (MAF 23.9%) [91]. A subsequent study [92] determined the MAF of rs1495741 in the Taiwanese population to be higher—almost 50%—and associated with a 14-fold increased risk of hepatotoxicity following antiTB treatment (isoniazid, pyrazinamide, and ethambutol) in patients with the AA genotype of rs1495741 [92].

This increase in risk is sufficient enough that the FDA has included genotypic information on the drug label for RIFATER® that reads "Slow acetylation may lead to higher blood levels of the drug and thus an in increase in toxic reactions" [93].

20.3.4 Sulfamethoxazole

Prophylactic sulfamethoxazole (SMX) is commonly prescribed for the prevention of *Pneumocystis jirovecii* pneumonia (PJP) in patients with HIV/AIDS. SMX is metabolized and excreted via NAT1 and NAT2 acetylation and is associated with a delayed hypersensitivity reaction [94]. The frequency of hypersensitivity reactions is notably greater in HIV-positive individuals (30–50%) when compared to non-HIV positive controls (3%) [95,96]. It has been suggested that the NAT2 slow acetylator phenotype is a risk factor for the development of SMX hypersensitivity reactions [97]. One study found a higher frequency of slow acetylators in SMX-hypersensitive HIV-positive patients [98,99], but other studies have disputed this association [100–102]. Further study suggests that the discrepancy between N-acetylator phenotype and genotype in HIV-positive patients results from the presence of rare alleles [103].

20.3.5 Irinotecan

The topoisomerase I inhibitor irinotecan is approved globally for the treatment of metastatic colorectal cancer. It is principally metabolized to its active metabolite SN38 by CYP3A and esterases. Detoxification of SN-38 occurs by glucuronidation facilitated by uridine diphosphate glucuronosyltransferase 1A1 (UGT1A1). Failure of this mechanism, and hence increased circulating unconjugated SN-38, puts patients at increased risk of severe neutropenia. The uridine diphosphate glucuronosyltransferases (UGTs) are a superfamily of enzymes primarily located in the endoplasmic reticulum of cells involved in the detoxification of exogenous compounds. The *UGT1A* gene loci is found at chromosome 2q37 and encodes all enzyme isoforms by alternative splicing of a unique exon 1 to a common domain that contains exons 2–5.

Among the most characterized polymorphisms in genes of the *UGT1A* family is a TA tandem repeat polymorphism in the *UGT1A1* promoter region. Alleles with 7 TA repeats, known as *UGT1A1*28,* have significantly lower transcription rates than does the wild-type allele (6 repeats) [104]. In one of the first examples of regulators applying pharmacogenomic biomarker information to dosing guidelines on a drug label, the FDA in 2005 advised that individuals who are homozygous for the *UGT1A1*28* allele are at increased risk of neutropenia following initiation of treatment with irinotecan, and dose reduction should be considered [105].

Numerous studies have confirmed the risk of neutropenia and/or diarrhea in carriers of the *UGT1A1*28* allele [106–112]. A meta-analysis of 821 patients suggested that, while *UGT1A1*28* homozygosity is indeed a risk factor for neutropenia, the risk is increased further in **28/*28* patients receiving high-dose irinotecan. Neutropenia risk in individuals receiving low-dose irinotecan was not significantly different among those homozygous for the wild-type allele, while high-dose **28/*28* genotype patients were at significantly greater risk (odds ratio = 27.8 [4.0–195]) [113]. Moreover, a meta-analysis of 16 trials in Caucasians confirmed that genotyping for the *UGT1A1*28* polymorphism before treatment can tailor therapy and reduce toxicity.

Although toxicity variants are greatest in *UGT1A1*28* homozygotes, those bearing the heterozygous variant are also at risk. Recently, a small phase II study of UGT 1A1 polymorphism-guided dose reduction in patients with metastatic colorectal cancer treated with irinotecan and 5′-deoxy-5-fluorouridine showed that dose reduction based on the **28* genotype is effective and safe. In wild-type **1/*1* individuals (22.7%) versus those carrying one variant **28* (16.7%) allele, no significant differences in response rate were identified and toxicities of grade 3 or higher including neutropenia (9.0% versus 0%, respectively) and diarrhea (13.6% versus 0%, respectively) were prevented. Importantly, another recent meta-analysis concluded that an individual's response to irinotecan is unlikely to be affected by *UGT1A1*28* status [114].

20.3.6 Combined PKPD

20.3.6.1 Warfarin

Approved in 1954, warfarin is an effective and widely prescribed anticoagulant for the prevention of thromboembolism. The International Normalized Ratio (INR) monitors the degree of anticoagulation in patients treated with it. Despite its effectiveness, warfarin's narrow therapeutic window and unpredictable interindividual variability poses a significant challenge to prescribers, particularly given the consequences of supratherapeutic INR with bleeding rates varying from 0.75% per year to as high as 10% in some studies [115,116].

Warfarin consists of a racemic mixture of *R* and *S* enantiomers and displays stereoselective metabolism; *S* warfarin is approximately three to five times more potent than *R* warfarin [117]. CYP2C9 is the main enzyme involved in the metabolic elimination of S warfarin. Warfarin's anticoagulant effect is achieved via binding of the 4-hydroxycoumarin ring to vitamin K epoxide reductase (VKOR), preventing reduction of vitamin K epoxide to dihydrovitamin K [118], the necessary co-factor required for the gamma-glutamyl carboxylation of vitamin K-dependent clotting factors [119,120]. As such, warfarin represents an excellent paradigm for how both PK and PD-related genetic variability can affect drug efficacy and safety.

Pharmacokinetics

CYP2C9 is the main enzyme involved in the metabolic elimination of S warfarin. Normal enzyme function (wild-type) is denoted *CYP2C9*1*, with the two most common allelic variants, *CYP2C9*2* (p.R144C) and *CYP2C9*3* (p.I359L) causing reductions in enzyme activity of 30% and 80%, respectively [121,122]. There are racial differences in the prevalence of CYP2C9 alleles. Approximately 1% of Caucasians are homozygous for the **2* variant; 0.4% are homozygous for **3* and heterozygous carriers for *CYP2C9*2*, *CYP2C9*3*, and *CYP2C9*2/*3*—representing about 22%, 15%, and 1.4% of the population, respectively. The *CYP2C9*2* and *CYP2C9*3* alleles are much less prevalent among Asians and African-Americans. Alteration in CYP2C9 enzymatic activity is associated with greater bleeding risk per given dose in patients treated with warfarin.

Pharmacodynamics

In the liver, warfarin targets the vitamin K epoxide reductase complex subunit 1 encoded by the *VKORC1* gene located on chromosome 16p11.2. Ten common non-coding *VKORC1* SNPs have been identified and grouped into five major haplotypes (H1, H2, H7, H8, and H9) with greater than 5% frequency in European Americans [123]. These can then be divided into two distinct evolutionarily distant groups, designated A (comprising H1 and H2) and B (comprising H7, H8, and H9); these groupings have been shown to significantly affect warfarin dose requirements, with genotype AA being the most sensitive and requiring the lowest dose and BB associated with the lowest enzyme activity and hence requiring the highest dose [123,124].

Regarding warfarin dose determination, in the last decade, substantial progress has been made in our understanding of how genotypes affecting CYP2C9 and *VKORC1* determine dose requirements [121–123,125–131]. When combined with environmental factors, genotypes affecting both CYP2C9 and *VKORC1* can predict around 50% of interindividual dose variation [132–134]. The initial 3-month period following warfarin commencement is considered "high-risk," with a 10-fold higher risk of bleeding events as compared to subsequent months [135]. Consequently, a combined genetic and clinical information approach could be adopted to guide the initiation of warfarin therapy during this period.

Warfarin labeling has undergone important revisions in the last decade. In 2007 and again in 2010, the FDA revised the drug label to incorporate into dosing decisions consideration of genotypic variation in *CYP2C9* and *VKORC1*. This led to the development of pharmacogenetic prediction algorithms [132,136,137]. Data from individual groups can be limited by geographical restriction and small sample size. In the largest study to date, the International Warfarin Pharmacogenetics Consortium (IWPC) demonstrated that the pharmacogenetic algorithm was significantly superior to the clinical algorithm in predicting stable warfarin dose in subjects who required <21 mg weekly and >49 mg weekly [132]. The IWPC is made up of 22 research groups across 11 countries; its investigators pooled genotype and phenotype data from 5700 warfarin-treated patients to produce a geographically and ethnically diverse population for study. Numerous prospective randomized controlled trials performed in recent years comparing pharmacogenetic with clinical algorithms have consistently demonstrated the superiority of the former [138–142]. Indeed, Finkelman et al. [138] demonstrated that, although genetic tables (as recommended by the FDA) predict warfarin dose better than empiric dosing, formal pharmacogenetic algorithms are the most accurate.

Bleeding Risk

Genetic variation contributes to increased bleeding risk. In the U.S. Medco and the Mayo clinic showed that warfarin genotyping reduced hospitalization rates for bleeding and thromboembolism by 43% [139]. In those carrying *CYP2C9* variants, patients have been shown to be at increased risk of over-anticoagulation and bleeding during induction

[125,126]. A meta-analysis involving 2775 patients demonstrated that this bleeding risk is greater with variants compared to carriers of the wild-type [143]. Patients with the *CYP2C9*2* and *CYP2C9*3* alleles had lower mean daily warfarin dose requirements and a greater risk of bleeding. In a racially diverse cohort followed prospectively for two years after initiation of warfarin, the variant *CYP2C9* genotype was associated with an increased risk for major hemorrhage [144].

The *VKORC1*−1639A allele is associated with higher INR values and more time spent with an INR above the upper limit of the therapeutic range [124,137,145,146]. Despite this, and in contrast to CYP2C9 variants, the majority of studies involving *VKORC1* polymorphisms have not shown any significant association between genetic variation in *VKORC1* and bleeding risk [124,144,147]. A single case-control study demonstrated an association between carriers of at least one T allele and increased risk of bleeding (crude odds ratio = 1.7, 95% CI: 1.1–2.5) compared to individuals with the wild-type genotype [148].

The majority of pharmacogenetic studies of warfarin have been conducted in Caucasian populations; therefore, the ability to predict the risk of over-anticoagulation may not be consistent across race. There is also lower linkage disequilibrium and greater variation in the African-American genome compared with non-Africans [149]. Consequently, genetic variants that influence warfarin sensitivity in non-Caucasians may go undetected in most studies. Indeed, resequencing of the *VKORC1* gene for SNP discovery was initially limited to an entirely Caucasian population [123]. A recent GWAS study in African-American individuals identified a novel CYP2C polymorphism that exerted a clinically important effect independent of the *CYP2C9*2* and **3* polymorphisms; this should also be incorporated into future genetic dosing algorithms [150].

Whether pharmacogenetic testing for warfarin is cost-effective remains to be seen. Ongoing prospective, randomized trials are evaluating its clinical utility and will provide important data to inform cost-effectiveness analyses. In particular, cost-effectiveness may vary by population, which could mean that testing will be restricted to those at greatest risk of adverse outcomes of supratherapeutic INR—the elderly, for example.

20.4 TYPE B ADVERSE DRUG REACTIONS

20.4.1 Malignant Hyperthermia and Anesthesia

Malignant hyperthermia (MH) is a potentially fatal hypermetabolic reaction of skeletal muscle that can be triggered by the administration of depolarizing muscle relaxants and/or halogenated anesthetics (e.g., halothane, enflurane, isoflurane). A MH reaction is characterized by dysregulated myoplasmic calcium (Ca^{2+}) leading to tachycardia, muscle rigidity, rhabdomyolysis, increased oxidative/nitrosative stress, mitochondrial damage, acidosis, hypercapnia, renal failure, and ultimately death [151]. MH was originally described in 1960 in a case report of a man who had experienced the deaths of ten close family members due to anesthetic complications [152]. The prevalence of MH has been estimated at 1.1 per 100,000 surgical discharges where there was indication for anesthesia [153]. MH occurs more frequently in men and younger patients (median age 22) [153–155]. MH-related mortality was estimated at 0.0082 per 100,000 surgical inpatients in a U.S. population [153].

MH is associated with a high morbidity rate of 35%, including coma, cardiac dysfunction, and renal dysfunction, which emphasizes the importance of improved treatments and prevention in future [154]. Dantrolene is the treatment of choice for MH and has reduced mortality from 80% in the 1960s to less than 10% today. Its pharmacological action is not completely understood, but inhibition of the ryanodine receptor, the major Ca^{2+} release channel of skeletal muscle sarcoplasmic reticulum, is believed to be fundamental in decreasing intracellular calcium concentration [156].

Susceptibility to MH is inherited as an autosomal dominant trait with variable expression and incomplete penetrance. Its exact pathophysiological mechanism is not completely understood, but uncontrolled release of Ca^{2+} from skeletal muscle sarcoplasmic reticulum disturbs the excitation-contraction coupling (ECC) of normal muscle contraction, leading to hypercontracture and hypermetabolism [157]. The cause of dysfunctional Ca^{2+} regulation in MH is, in most cases, secondary to defects in the ryanodine receptor type 1 (RYR1) and/or the dihydropyridine-sensitive L-type voltage-dependent calcium channel (DHPR).

20.4.1.1 Diagnosis

Individuals with MH are phenotypically normal and will only manifest symptoms of MH when exposed to triggering anesthetic agents. In patients who have had a suspected episode of MH, confirmation of the diagnosis is crucial because the condition is dominantly inherited and could affect other family members. Currently, diagnosis is based on a combination of muscle contracture testing and molecular genetic screening [158].

Muscle Contracture Testing

The "gold standard" for diagnosis of MH is the muscle contracture test, of which two versions exist: the *in vitro* contracture test (IVCT) devised by the European Malignant Hyperthermia group (EMHG); and the caffeine-halothane contracture test (CHCT) devised by the North American Malignant Hyperthermia Group (NAMHG) [159,160]. The two tests differ in methodology but essentially involve a muscle tissue biopsy from the affected individual, exposure of the muscle fibers to halothane and caffeine, and then observation for a positive contractile response (Table 20.7) [158].

Despite these slight differences, the EMHG IVCT has a sensitivity of 99% and a specificity of 93.6%, while the NAMHG CHCT has a sensitivity of 97–98% and a specificity of 78–80% [161,162]. Both versions have very high sensitivity values because MH is a potentially fatal adverse reaction and anesthesiologists are more likely to accept false-positive results compared with false-negative results. Although the diagnostic gold standard, the muscle contracture test is invasive and expensive, and can only be done in specialized centers. Therefore, alternatives such as molecular genetic testing have been introduced.

Genotype Testing

Genetic testing for MH was introduced into clinical practice in 2005 [163]. More than 300 mutations have been identified with *RYR1*, although only 30 of these have sufficient evidence to be considered causal for MH [164]. Approximately 30% of patients with MH carry one of the 30 variants in *RYR1* that are known to be causal, while up to 70% of MH patients carry a *RYR1* DNA variant, whether causal or of unknown significance [165,166]. To determine if a variant is causally associated, it must be characterized through genetic and functional studies while satisfying all the requirements set out by the (EMHG) (http://www.emhg.org/documents/) (Table 20.8).

The other genetic association reported with MH is in *CACNA1S*, which encodes the alpha-1 subunit of the skeletal muscle DHPR and accounts for approximately 1% of cases of MH [167].

Genetic diagnosis of susceptibility to MH by analysis of *RYR1* and *CACNA1S* is difficult because of significant genetic heterogeneity and discordance between genotype and phenotype. No single dominant mutation is causal for MH, and the *RYR1* and *CACNA1S* genes are complex, comprising 106 exons and 46 exons, respectively [168,169]. The A1352G substitution in *RYR1* is associated with MH in Caucasian patients, but in individuals of African descent the polymorphic change appears to be neutral, suggesting that DNA mutation profiles may differ between human populations [170]. Adding further complexity is the varying prevalence of specific mutations in different populations. In the United Kingdom, the most common mutation is G2434R, whereas in Germany and France R163C is more commonly reported [171]. These findings highlight the importance of not only screening for mutations among different ethnicities, but also screening of large numbers of control patients to exclude nonsignificant mutations.

Currently, the major limitation of genetic screening for *RYR1* in the diagnosis of MH is that a negative result cannot exclude MH without further muscle contracture testing. Other limitations include low sensitivity (25%) and considerable inter- and intraindividual variation in phenotypic expression [172]. However, genetic testing does have a role in screening family members of patients who have been definitively clinically diagnosed with MH and where a causative mutation has been identified. This mutation is then screened for in-family members; if they are positive, they are regarded as MH susceptible without muscle contracture testing [158,173].

Another limitation of genetic testing is the significant variation in phenotype expression and penetrance in individuals who carry genetic mutations associated with MH, although mutations that occur in highly conserved regions of *RYR1* seem to be associated with a more severe MH phenotype [163,174]. The role of epigenetic factors and the effects of multiple mutations in a single *RYR1* gene remain unclear [175,176]. Until our understanding of the relationship between genotype and phenotype improves, the goal of genetic testing alone to diagnose MH in clinical practice remains distant.

20.4.2 Statins and Skeletal Muscle Toxicity

The 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA reductase) inhibitors (statins) prevent the

TABLE 20.7 Diagnostic Criteria for European and American Muscle Contracture Testing

	In Vitro Contracture Test (IVCT)	Caffeine-Halothane Contracture Test (CHCT)
MH susceptible	Positive response to both halothane and caffeine	Positive result for either halothane or caffeine
MH normal	Negative response to both halothane and caffeine	Negative response to both halothane and caffeine
MH equivocal	Positive response to either halothane or caffeine	Not applicable

TABLE 20.8 EMHG Criteria for Gene Mutations to be Used in Genetic Testing of MH Susceptibility

Genetic characterization	• Full description at DNA and protein level, considering evolutionary conservation and change in charge, polarity, or structure introduced by amino acid replacement • Co-segregation of mutation with disease in at least 2 pedigrees • Absence of sequence change from 100 control samples for exclusion of polymorphism
Functional characterization	• Recombinant *in vitro* expression on a defined background (e.g., expression of mutated RYR1 cDNA construct in cells that do not express RYR1) and measurement of Ca^{2+} release in response to RYR1-agonists • Alternatively, myotubes of RYR1-knockout mice can be used as recipients for expression of cDNA constructs

Assays of RYR1 function in *ex vivo* tissues	• Calcium measurement ligand-binding studies performed on tissues from MH patients with characterized RYR1 mutations:
	  • Myotubes
	  • Microsomal SR preparations from muscle biopsies
	  • Lymphoblasts
	• To avoid interference of other genetic factors, all assays based on cells taken from patients should be performed on samples from at least 2 independent patients with same mutation

biosynthesis of cholesterol, thereby lowering low-density lipoprotein cholesterol (LDL-C) and reducing cardiovascular events. Statins are generally well tolerated, and the incidence of ADRs is rare. However, safety concerns arose following the worldwide withdrawal in 2001 of cerivastatin because of an unacceptably high incidence of rhabdomyolysis. In general, muscle toxicity with statins ranges from myalgia (muscle pain without any elevation in creatine kinase (CK)) to rhabdomyolysis (severe muscle damage where the CK is >10 times the upper limit of normal) [177]. A systematic review of 20 RCTs in statin-treated patients versus placebo defined the incidence of rhabdomyolysis as 3 per 100,000 person-years in those taking simvastatin, lovastatin, atorvastatin, pravastatin, or fluvastatin [178]. Muscle toxicity inevitably affects compliance; 25–50% of patients with coronary artery disease (CAD) are noncompliant with a prescribed drug at 12 months [179]. Thus, although the incidence of rhabdomyolysis is low, other muscle symptoms, in particular muscle pain without a rise in CPK, may be more common, and likely to lead to drug discontinuations.

Although the exact cellular mechanism underlying the myopathic effect of statins has not been fully elucidated, it is clear that it is dose-dependent [180] and increases with increasing blood concentrations [181]. SLCO1B1 is an organic anion transporting polypeptide exclusively localized to the hepatic basolateral membrane; it is a key mechanism for hepatic uptake of statins. Functional variation of SLCO1B1 has been implicated in statin-induced myopathy and comprises 5 haplotypes—*1a wild-type, *1b usual transport activity, *5 (c.521T>C alone), *15 (combined c.521T>C and c.388A>G), and *17—all with reduced transport activity. Reduced transport ultimately leads to increased circulating statin concentrations and, indeed, increased AUCs have been seen in carriers of at least one copy of the *5 or *15 allele taking atorvastatin [182], rosuvastatin [182], pravastatin [183], and simvastatin [184].

In particular, variation involving the *5 allele (rs4149056), has been associated with higher rates of statin intolerance [185–187] and high risk of developing statin myopathy as determined by GWAS [188]. The risk of statin-induced myopathy is additive, dependent on the number of *5 alleles carried. The incidence in those with no, one, or two *5 alleles has been shown to be 20%, 27%, and 50%, respectively [187]. The greatest risk appears to occur with simvastatin. Some studies failed to show a comparable risk with atorvastatin [188,189], pravastatin [187], or rosuvastatin [190], but the small sample size of these studies restricts the validity of the findings. A recent large real-world study of 288 Caucasian individuals treated with atorvastatin or rosuvastatin identified a 45-fold interindividual variability in plasma levels in those taking the same dose, which is almost entirely attributable to genetic variation. Individuals were genotyped for common uptake and efflux transporter polymorphisms along with drug metabolism polymorphisms. Measurement of 4B-hydroxycholesterol concentration served as a marker of CYP3A4; lathosterol concentration, as a measure of statin-mediated inhibition of de novo cholesterol synthesis. Ninety percent of rosuvastatin variability is attributable to the reduced-function transporter polymorphisms in SLCO1B1 and ABCG2; for atorvastatin, this is divided between SLCO1B1 and CYP3A4. In addition, retrospective analysis of 579 patients taking statins revealed that almost 50% of those taking the highest doses have concentrations above the 90th percentile, suggesting the appearance of a high-risk group. The authors propose a dosing algorithm for statin prescribing that incorporates genomic and clinical variables in order to reduce the risk of statin myopathy [191].

More recently, a breakthrough study reported that levels of expression quantitative loci (eQTL) varied between simvastatin-exposed and non-simvastatin-exposed controls *in vitro*. The most significant eQTL with simvastatin exposure was associated with the gene glycine amidinotransferase (*GATM*), which encodes the enzyme for the rate-limiting step of creatine synthesis. Variation in this locus was demonstrated to affect cholesterol synthesis and metabolism and was significantly associated with statin-induced myopathy. The authors hypothesize that GATM may represent the link between statin-induced myopathy and cellular cholesterol homeostasis and energy metabolism [192].

20.4.3 Drug-Induced Skin Injury

Drug-induced skin injury (DISI) is common and has diverse clinical manifestations ranging from mild inflammation, such as maculopapular exanthema (MPE), to severe life-threatening reactions that include Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN), acute generalized exanthematous pustulosis (AGEP), and hypersensitivity syndrome (HSS) (Table 20.9) [193]. DISI reactions are estimated to account for 5–15% of all ADRs [194]. The strongest genetic associations with ADRs have been reported with human leukocyte antigens (HLA) (Table 20.10). HLA alleles are encoded by the major histocompatibility complex (MHC) on chromosome 6 and play an integral role in immune response to antigens. It has been proposed that specific HLA molecules present drug or drug metabolites to T cells, which trigger the immune system and lead to a hypersensitive reaction [195]. The exact mechanism by which this occurs is not definitively known and may vary with different drugs. Several hypotheses have been put forward that are outside the scope of this chapter. Readers are referred to excellent reviews for further information [196,197].

20.4.3.1 *Abacavir Hypersensitivity*

Abacavir is a nucleoside reverse transcriptase inhibitor prescribed for the treatment of human immunodeficiency virus (HIV) infection [198]. Approximately 5% of patients experience abacavir hypersensitivity syndrome (AHS), with clinical features similar to HSS. It usually presents within 6 weeks of initiation of abacavir, which necessitates immediate cessation of therapy. Re-exposure of these patients to abacavir leads to rapid reappearance of symptoms and the potential to induce more severe symptoms [199]. The clinical features of AHS and the discovery of CD8+ T-cell response in skin biopsies of abacavir patch test-positive patients have provided evidence for an immune-based etiology in AHS [200]. Genetic susceptibility is suggested by familial disposition to AHS, reduced risk in black ethnicities, and the observation that only a minority of patients develop hypersensitivity within six weeks, while those who do not remain at low risk despite long-term therapy [201,202].

AHS was first reported to be significantly associated with *HLA-B*57:01* in an Australian cohort [203]. In this cohort, 14 out of 18 (78%) AHS patients were positive for HLA-*B*57:01* compared with 4/167 (2%) of abacavir-tolerant cases (OR = 117, $p < 0.0001$, 95% CI 29–481). Screening

TABLE 20.9 Phenotype Definitions in Drug-Induced Skin Injury

Phenotype	Clinical Features
Stevens-Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN)	Skin detachment: 1–10% in SJS 10–30% in overlap syndrome >30% in TEN Severe erosions of mucous membranes—may be hemorrhagic Systemic involvement: Fever Deranged liver function tests Lymphopenia Intestinal manifestations Pulmonary manifestations
Hypersensitivity Syndrome (HSS); also known as drug-induced hypersensitivity syndrome (DIHS); drug reaction with eosinophilia and systemic symptoms (DRESS); drug-induced delayed multiorgan hypersensitivity syndrome	Fever (typically >38°C) Exanthema Lymphadenopathy Internal organ involvement Eosinophilia Atypical lymphocytes Lymphopenia Thrombocytopenia
Acute generalized exanthematous pustulosis (AGEP)	Acute widespread edematous erythema followed by sterile pustular eruption Pustules initially located in neck, groin, and axillae; later become widely disseminated Fever (>38°C) Neutrophilia +/– mild eosinophilia

TABLE 20.10 Drug-Induced Skin Injury and HLA Allele Associations

Drug	Class of Drug	HLA Association	Ethnicity	Phenotypes
Abacavir	Antiretroviral	*B*57:01*	Caucasians	HSS
			African-Americans	
Carbamazepine	Anticonvulsant	*B*15:02*	Han Chinese	SJS/TEN
			Thai	
			Malaysian	
			Indian	
		*A*31:01*	Japanese	MPE
			Caucasian	HSS
			Korean	SJS/TEN
			Han Chinese	
		*A*02:01*	Han Chinese	MPE
		*DRB1*14:05*		
		*B*58:01*		
Oxcarbazepine	Anticonvulsant	*B*15:02*	Han Chinese	MPE
		*B*13:02*		
Allopurinol	Xanthine oxidase inhibitor	*B*58:01*	Han Chinese	HSS
			Thai	SJS/TEN
			Japanese	
			Malay	
Phenytoin	Anticonvulsant	*B*15:02*	Han Chinese	SJS/TEN
			Thai	
Lamotrigine	Anticonvulsant	*B*58:01*	Caucasian	HSS
				SJS/TEN
		*A*68:01*		
		*DRB1*13:01*		
		*A*30:01*	Han Chinese	MPE
		*B*13:02*		
		*A*33:03*		
Nevirapine	Antiretroviral	*DRB1*01*	Caucasian	
		Cw8, B14	Caucasian	MPE
		Cw8	Japanese	HSS
		*B*35:05*	Thai	SJS/TEN
			Indian	
		*Cw*04*	Han Chinese	
			Thai	
Sulfamethoxazole	Antibiotic	*B*38, A29, B12, DR7*	Caucasian	SJS/TEN
Aminopenicillins	Antibiotic	*A2, Drw52*	Caucasian	HSS
Co-trimoxazole	Antibiotic	*A30, B13, Cw6*	Turkish	Fixed drug eruption

Source: Adapted from Yip VLM, Pirmohamed M. Genetics of adverse drug reactions. Clin Rev Allergy Immunol, in press.

for *HLA-B*57:01* was reported to have a positive predictive value (PPV) of 100% and a negative predictive value (NPV) of 97%. The association was subsequently replicated in populations from North America and the United Kingdom [204,205].

Once the association between *HLA-B*57:01* and AHS was confirmed, subsequent research efforts focused on translating the association into a clinically applicable test for identification of susceptible individuals. Simple tests for *HLA-B*57:01* were developed and prospective screening has significantly reduced the incidence of AHS from 8% to less than 1% across multiple populations and ethnicities including the United Kingdom, France, and Australia [206–210]. The PREDICT-1 prospective controlled trial randomized 1956 patients from 19 countries into *HLA-B*57:01* screening or a standard-care approach of abacavir therapy without screening. The results were that the incidence of AHS in the prospective screening group was 3.4% compared with 7.8% in the control group ($p<0.001$), confirming the clinical utility of pretreatment *HLA-B*57:01* testing in abacavir treatment [211].

Three studies across different populations addressed the issue of cost-effectiveness of screening for *HLA-B*57:01* prior to initiation of abacavir therapy [205,212,213]. *HLA-B*57:01* screening before abacavir therapy was concluded to be cost-effective but dependent on the costs and efficacy of alternative therapies, such as tenofovir-containing antiretroviral regimens. However, it is important to bear in mind that cost and efficacy are not the only factors that influence the choice of antiretroviral therapy in HIV. Clinician preferences, the low incidence of other side effects (e.g., lipodystrophy), and other factors such as blood–brain barrier penetrance favor abacavir therapy. More important, HIV mutation and resistance to antiretrovirals mean that clinicians require the availability of all anti-HIV treatments, of which abacavir is one of a finite number [205]. Screening for *HLA-B*57:01* is now the most frequently used pharmacogenetic test in clinical practice.

20.4.3.2 Carbamazepine Hypersensitivity

Carbamazepine (CBZ) is an important drug that is used in the treatment of epilepsy, trigeminal neuralgia, and bipolar disorder [214–216]. Cutaneous ADRs have been reported with CBZ in up to 10% of patients [214]. These reactions can range from MPE to severe HSS or life-threatening SJS/TEN [217]. The annual incidence of SJS has been estimated to be 1–6 cases per million, and TEN has an incidence of approximately 1–2 cases per million. Of cases of TEN, 80–95% are believed to be drug-induced while 50–80% of cases of SJS are attributable to drugs [218]. The drugs most commonly associated with SJS/TEN are nevirapine, anticonvulsants (carbamazepine and lamotrigine), antibiotics (sulphonamides and aminopenicillins), allopurinol, and oxicam nonsteroidal anti-inflammatory agents [219]. Other causes of SJS/TEN include vaccination, mycoplasma pneumonia infection, dengue virus, and intravenous administration of contrast agents [220]. TEN is associated with significant mortality, up to 40%, and long-term morbidity, including blindness and scarring [217].

Researchers from Taiwan were the first to report a very strong association between *HLA-B*15:02* and CBZ-induced SJS among a Han Chinese population ($\text{OR}=2504$, $p=3.1\times10^{-27}$, 95% CI 126–49,522) [221]. This association was replicated in Han Chinese patients from China and Asian populations including Thai, Malay, and Indian [222–228]. However, it should be noted that European, Japanese, and Korean patients with CBZ-induced SJS/TEN did not demonstrate this association. Furthermore, *HLA-B*15:02* was not associated with other forms of CBZ-induced cutaneous ADR (HSS or MPE). Therefore, the association between *HLA-B*15:02* and CBZ-induced ADR is specific to both phenotype and ethnicity.

The incidence rate of SJS/TEN in some Southeast Asian countries, such as Taiwan, Malaysia, and Thailand, is 10-fold higher than the rates observed in Europe, the United States, or Japan [229]. This discrepancy may be explained by the differing background allele frequencies among different populations: the prevalence of *HLA-B*15:02* is highest among Asian populations (0.057–0.145 in Han Chinese, 0.085–0.275 in Thais, and 0.12–0.157 in Malays) compared with European (0.01–0.02), Japanese (0.002), and Korean (0.004) populations [229]. A systematic review demonstrated that carriage of the *HLA-B*1502* allele was associated with a 100-fold risk of developing SJS/TEN with CBZ [230]. The number needed to test in order to avoid an episode of CBZ-induced SJS/TEN was estimated to be 461. The clinical utility of preprescription studies was demonstrated by a prospective cohort study conducted in Taiwan [231], in which 4877 patients were prospectively genotyped for *HLA-B*15:02* and patients who were positive were prescribed alternative therapy to CBZ. The percentage of patients positive for *HLA-B*15:02* was 7.7. Based on the historic incidence of CBZ-induced SJS/TEN, 10 cases of SJS would have been expected, but in the Taiwanese study no patients developed either condition.

Two cost-effectiveness studies, in Thailand and Singapore, concluded that screening for *HLA-B*15:02* could be cost-effective in certain ethnicities, such as Singaporean Chinese and Malays, but would depend on population frequency of *HLA-B*15:02*, duration of treatment, PPV, and costs of alternative drugs and of the genotyping test itself [232,233]. On the basis of these data, the FDA and the European Medicines Agency now include warnings on the drug label advising clinicians of the need to genotype patients from certain areas of Asia prior to commencing therapy with CBZ [234].

In Japanese and Korean patients *HLA-B*15:11* has been associated with CBZ-induced SJS/TEN [235,236]. However, this association needs to replicated in larger numbers of patients and other populations in Europe before it can be considered as a screening test for CBZ-induced cutaneous ADRs.

Two GWAS, in Caucasian and Japanese patients, independently reported associations between *HLA-A*31:01* and all phenotypes of CBZ hypersensitivity, unlike *HLA-B*15:02*, which was phenotype specific for SJS/TEN [237,238]. This association was also demonstrated in Han Chinese and Korean patients [236,239]. Carriage of *HLA-A*31:01* was associated with a 9-fold increased risk of CBZ cutaneous ADR, and the number needed to test to avoid one case was between 47 and 67 depending on ethnicity [230].

20.4.3.3 Allopurinol Hypersensitivity

Allopurinol is a xanthine oxidase inhibitor indicated for the treatment of gout and hyperuricemia [240]. A case-control study in Han Chinese patients first reported an association between *HLA-B*58:01* and allopurinol-induced SJS/TEN and HSS [241]. In this study, all 51 patients who had developed allopurinol-induced SJS/TEN or HSS were positive for *HLA-B*58:01* compared with 20 out of 135 allopurinol-tolerant controls (15%) and 19 out of 93 (20%) in the general population ($OR = 580$, $p_c = 4.7 \times 10^{-24}$, 95% CI 34–9781). The association was replicated in patients from other Southeast Asian countries, including Thai, Japanese, and Korean [235,242–245]. In contrast to the association with CBZ, *HLA-B*58:01* was also identified as a potential biomarker for allopurinol-induced SJS/TEN and HSS in European populations [246]. A genome-wide association study (GWAS) in 14 Japanese patients with allopurinol-induced SJS/TEN identified the single nucleotide polymorphism (SNP) rs9263726 in *PSORS1C1* to be strongly associated with these conditions ($OR = 61.2$, $p = 3.64 \times 10^{-8}$, 95% CI 18.4–203.5). This SNP was in absolute linkage disequilibrium with *HLA-B*58:01* and may represent a simple and useful biomarker for screening prior to allopurinol therapy in Japanese patients. More recently, an association has been reported between *HLA-B*5801* and allopurinol-induced MPE in a Han Chinese population from China [242]. In this cohort, all 22 patients with allopurinol-induced MPE possessed *HLA-B*5801* compared with 7 out of 63 allopurinol-tolerant controls (11.1%) ($OR = 339$, $p = 9.2 \times 10^{-14}$, 95% CI 18.6–6186.4). However, the association between *HLA-B*58:01* and MPE needs to be replicated in other populations. The clinical utility of preprescription genotyping for *HLA-B*58:01* has not been determined and requires further research before it can be employed in routine clinical practice.

20.4.3.4 Maculopapular Exanthema Reactions

MPE is the most common cutaneous manifestation of ADR and accounts for 90–95% of skin reactions [247]. It presents as a diffuse cutaneous erythema with areas of skin elevation that usually appear on the trunk and spread to the extremities in a symmetrical pattern. Pruritus, low-grade fever, and eosinophilia are frequently present. The reaction usually presents in the first 7–14 days on first exposure but reappears more rapidly, 2–3 days, on re-exposure. MPE is usually self-limiting and clears 1–2 weeks following withdrawal of the causative agent [248]. In some cases, it can progress to more severe manifestations of DISI, but predicting which cases will progress is not possible at present. The drugs associated most frequently with MPE are penicillins, sulphonamides (especially co-trimoxazole), and NSAIDs [247]. Genetic associations to MPE have been reported; these are caused by antituberculous treatment and antiepileptic therapies.

Antituberculosis Medications

Treatment for TB is commonly associated with MPE reactions and has an estimated incidence of 0.1–5.7% among patients receiving first-line agents: isoniazid, rifampicin, ethambutol, and pyrazinamide [249]. In some cases, patients need to discontinue medication or change to other less effective regimens, leading to delay in TB treatment [250]. A recent study reported that polymorphisms in *CYP2C19* and *CYP2C9* were significantly associated with an increased risk of developing antiTB drug- (ATD-) induced MPE [251]. In this study, 62 patients with first-line ATD-MPE were compared against 159 patients without any adverse events to first-line ATD. The SNP −1565 C>T in *CYP2C9* showed a significant association with ATD-induced MPE, with fewer patients carrying the minor alleles (CT or TT) in cases compared with controls (4.9% versus 18.5%, $p_c = 0.022$, $OR = 0.23$, 95% CI 0.07–0.78). The SNP −636 G>A in *CYP2C19* (*CYP2C19*3*) was also significantly associated with susceptibility to ATD-induced MPE. Only 6.6% of patients with ATD-induced MPE carried the minor allele-containing genotypes (GA or AA) compared with 19.1% of ATD-tolerant controls ($p_c = 0.042$, $OR = 0.27$, 95% CI 0.09–0.82). One haplotype of *CYP2C9* and *CYP2C19* containing both significant SNPs (−1418 C>T-636 G>A-1565 C>T_−1188 C>T) was significantly associated with development of ATD-induced MPE ($p = 0.012$, $OR = 0.13$, 95% CI 0.03–0.57). No significant associations were reported between polymorphisms in *NAT2* and *CYP2E1* and ATD-induced MPE.

CYP2C9 and CYP2C19 are involved in the metabolism of both rifampicin and isoniazid [252]. Carriage of the *CYP2C9* polymorphism −1565 C>T appears to be protective against ATD-induced MPE. The functional significance of this polymorphism is currently not known, but has been hypothesized to affect either enzyme expression or enzyme activity, thereby affecting drug metabolite synthesis [251]. The polymorphism also demonstrates high linkage disequilibrium with the *CYP2C19*3* polymorphism, and it may be that the association between *CYP2C9* −1565C>T is actually related to altered CYP2C19 activity. *CYP2C19*3* is a poor metabolizer phenotype with a frequency of 0.06–0.10 among Asian populations [253]. Carriage of *CYP2C19*3* was protective for ATD-induced MPE, suggesting that modified metabolism and metabolite formation may be responsible for susceptibility to ATD-induced MPE [251].

20.4.3.5 Antiepileptics

Antiepileptic drugs (AEDs) are one of the most common culprits in triggering DISI. Carbamazepine (CBZ), lamotrigine (LTG), and oxcarbazepine (OXC) have all been associated with MPE [254,255]. When a patient develops a cutaneous ADR with AED therapy, early withdrawal of treatment is necessary because late withdrawal has been associated with negative outcome in patients who progress to more severe forms of DISI [256]. Several associations have been reported between AED-induced MPE and HLA [237,239,257–259].

A significant association between CBZ-induced MPE and *HLA-A*31:01* was originally reported in Han Chinese patients from Taiwan [239]. In this cohort, 18 patients were diagnosed with CBZ-induced MPE and 6 carried the *HLA-A*31:01* allele; this compared with only 6 in the CBZ-tolerant group ($p_c = 2.2 \times 10^{-3}$, OR = 17.5, 95% CI = 4.6–66.5). This association was replicated in European patients, where 22% with CBZ-induced MPE were positive for the *HLA-A*31:01* allele compared with only 4% who were CBZ tolerant ($p = 8.0 \times 10^{-7}$, OR = 8.33, 95% CI 3.59–19.36) [237]. More recently, a larger study of Han Chinese patients identified four new risk alleles for the development of CBZ-induced MPE [258]. Two alleles, *HLA-A*02:01* and *HLA-DRB1*14:05*, were found at significantly increased frequency among CBZ-induced MPE patients compared with tolerant controls (p = 0.033 and 0.003, respectively). *HLA-B*58:01* was present in 4% of CBZ-MPE patients compared with 13% of CBZ-tolerant controls, and the frequency of *HLA-DRB1*03:01* in CBZ-MPE patients was 4% compared with 13% in tolerant controls. These studies suggest that the *HLA-A*31:01*, *HLA-A*02:01*, and *HLA-DRB1*14:05* alleles increase the risk of CBZ-induced MPE, whereas *HLA-B*58:01* and *HLA-DRB1*03:01* protect against it.

LTG is an anticonvulsant with an aromatic ring structure similar to that of CBZ. It is associated with cutaneous ADRs with a reported frequency of approximately 4.8% [260]. Three alleles have been associated with LTG-induced MPE: *HLA-A*30:01*, *HLA-B*13:02*, and *HLA-A*33:03* [258]. The frequency of *HLA-A*30:01* and *HLA-B*13:02* was significantly increased in patients with LTG-induced MPE compared with LTG-tolerant controls. *HLA-A*33:03* appeared to have a protective effect, with 3% of LTG-MPE patients carrying the allele compared with 11.4% of LTG-tolerant controls (p = 0.048). However, it is important to note that two other studies failed to detect any significant associations between HLA alleles and LTG-induced cADR [261,262].

OXC is an analog of CBZ with antiseizure efficacy equivalent to that of CBZ. It is considered to be safer than CBZ because of its metabolism via alternative metabolic pathways, and has therefore been recommended as an alternative to CBZ [255,263]. However, recent studies have reported the incidence of OXC-induced cutaneous ADR to be similar to that resulting from CBZ, at 8%–9% [264,265]. *HLA-B*15:02* was reported to be significantly associated with OXC-induced MPE in a population of Han Chinese patients [257]. In this study, 44.4% of patients with OXC-induced MPE possessed the *HLA-B*15:02* allele compared with 11.1% in the OXC-tolerant group and 8.3% among population controls. However, only the association between OXC-induced MPE patients and population controls reached significance (OR = 8.8, p = 0.011, 95% CI 1.853–41.790). A second study of Han Chinese patients failed to replicate this association, with only 3.6% of patients with OXC-induced MPE positive for *HLA-B*15:02* [259]. This study did report that *HLA-B*13:02*, with a frequency of 14.3% among OXC-MPE patients compared with 2.1% in the general population, is a significant risk allele for OXC-induced MPE (OR = 7.83, p = 0.001, 95% CI 2.32–26.41). It must be noted that both of these studies included small numbers of patients and require further validation in much larger cohorts.

20.4.4 Drug-Induced Liver Injury

Drug-induced liver injury (DILI) is a rare, but potentially fatal, adverse hepatic reaction to medication. Because it is usually unexpected on the basis of the pharmacological action of the drug administered, it is distinct from DILI secondary to drug overdose, such as acetaminophen, where the risk and severity of injury generally increases with the dose taken [266]. DILI has an estimated incidence of 1–10 per 100,000 in patients exposed and accounts for up to 15% of acute liver failure cases in Europe and the United States. It is the single most common adverse effect leading to drug withdrawal or FDA refusals to approve [267]. Drugs that cause DILI are chemically diverse and belong to a multitude of therapeutic classes. In the majority of treated patients, these drugs are safe over a wide range of dosages, but a small subset of patients experience severe ADRs (e.g., DILI). Because of this unpredictability and low incidence of

TABLE 20.11 Drug-Induced Liver Injury and Reported Genetic Associations

Drug	Class of Drug	Genetic Association	Ethnicity
Amoxicillin-Clavulanate	Antibiotic	*DRB1*15:01*	Caucasians
		*A*02:01*	Spanish
		*B*18:01*	
Flucloxacillin	Antibiotic	*B*57:01*	Caucasian
Ximelagatran	Oral direct thrombin inhibitor	*DRB1*07:01*	Various
		*DQA1*02:01*	
Lumiracoxib	COX-II inhibitor	*DRB1*15:01*	Various
Nevirapine	Antiretroviral	*DRB1*01:01*	Caucasian
			Australian
Lapatinib	Tyrosine kinase inhibitor	*DQA1*02:01*	Various
Ticlopidine	Antiplatelet	*A*33:03*	Japanese
		*A*33:03 & CYP2B6 *1H/*1J*	Japanese
Antituberculosis	Antibiotic	*DQB1*02:01*	Indian
Troglitazone	Antidiabetic	*GSTM1-GSTT1 double null*	Japanese
Valproate	Anticonvulsant	*POLG*	Caucasian

Source: Adapted from Yip VLM, Pirmohamed M. Genetics of Adverse Drug Reactions. Clin Rev Allergy Immunol, in press.

events, a drug's hepatotoxic potential may be realized only after it has obtained market approval and large numbers of patients have been exposed [268–270].

Diagnosis of DILI is difficult because of variability in presentation, heterogeneity of hepatotoxic medications, and lack of understanding of the pathophysiology of liver injury. A case definition and diagnostic algorithm for DILI has been published that will improve diagnostic accuracy and identification of culprit medications [266]. The mechanism of drug-induced hepatotoxicity is a complicated multistep process that involves direct injury to the hepatocytes and activation of inflammatory pathways. It is thought that initial events in DILI are drug specific, and that the drug and/or its metabolites directly cause cell stress, trigger immune response, and/or impair mitochondrial function. Subsequent "downstream" steps are less specific, and common to all drugs, representing a balance between protective and toxic pathways. In susceptible individuals, apoptotic or necrotic cell death is initiated when "mitochondrial permeability transition" is activated as toxic pathways overcome protective mechanisms [271].

Risk factors for DILI susceptibility include genetic variation in drug metabolism pathways and immune response [272]. As with DISI, several associations have been reported between HLA and DILI (Table 20.11). This is a little surprising because the clinical features of DILI with some drugs (e.g., lumiracoxib, ximelagatran) do not suggest an immune component [273,274]. Despite the strong associations reported by some studies, none of these tests have been translated into clinical practice.

20.4.4.1 Amoxicillin-Clavulanate

Amoxicillin-clavulanate (AC) is one of the most commonly prescribed antibacterials worldwide. It is an important cause of DILI, accounting for approximately 10–13% of DILI-related hospitalizations [275,276]. The clinical presentation of AC-induced DILI is varied, with most patients recovering after cessation of AC but with a minority developing acute liver failure leading to mortality or requiring transplantation [277,278]. An association between HLA and AC-DILI was first reported in a European population where the frequency of the haplotype *HLA-DRB1*15:01-DRB5*01:01-DQB1*06:02* was significantly higher in the group with AC-DILI compared with population controls (57.1% versus 11.7%, $p_c < 0.0002$) [279]. This association was replicated in two UK populations [280,281]. In the first cohort, 14 of 20 AC-DILI patients (70.0%) were positive for *HLA-DRB1*1501* compared with 27 of 134 controls (20.1%) ($OR = 9.3$, $p < 2.5 \times 10^{-6}$). All patients positive for *HLA-DRB1*1501* also possessed the *HLA-DRB1*15:01-DRB5*01:01-DQB1*06:02* haplotype, as reported in the first AC-DILI study [280]. The third study confirmed the association between *HLA-DRB1*15:01* and AC-DILI (OR = 2.6, p = 0.002, 95% CI 1.44–4.68), but also identified a protective genotype, *HLA-DRB1*07* (OR = 0.266, $p_c = 0.0266$, 95% CI 0.11–0.65) [281].

It should be noted that *HLA-DRB1*15:01* was not significant in a Spanish cohort of patients with AC-DILI [282]. A GWAS study involving 201 patients of White European or U.S. origin with AC-DILI replicated the association between *HLA-DRB1*15:01* and AC-DILI (OR = 4.2, $p = 4.6 \times 10^{-6}$, 95% CI 2.7–6.6). Two novel risk HLA alleles were also identified: *HLA-A*02:01* (OR = 2.2, $p = 2.0 \times 10^{-6}$, 95% CI 1.6–3.2) in all patients, and *HLA-B*18:01* (OR = 4.0, p = 0.004, 95% CI 1.5–11) in Spanish patients only [283]. Carriers of both *HLA-DRB1*15:01* and *HLA-A*02:01* have approximately a 10-fold increased risk of AC-DILI. Assuming a 0.014% incidence of AC-DILI, the PPV and NPV for testing both alleles is 0.1% and 99.9%, respectively. This low PPV means that prospective genotyping is unlikely to be effective at identifying those at risk of DILI but may have clinical utility in strengthening AC-DILI diagnoses.

20.4.4.2 Flucloxacillin

Flucloxacillin is an antimicrobial agent used widely in Europe for treatment of methicillin-sensitive staphylococcal infections. Flucloxacillin use has been associated with a characteristic cholestatic hepatitis with an estimated incidence of 8.5 per 100,000 users [284]. A GWAS in 51 patients reported a strong association between *HLA-B*57:01* and flucloxacillin DILI (OR = 80.6, $p = 9.0 \times 10^{-19}$, 95% CI 22.8–284.9) [285]. A replication cohort of 23 further cases of flucloxacillin DILI confirmed the association (OR = 100.0, $p = 6.62 \times 10^{-13}$, 95% CI 20.6–485.8). Therefore, patients positive for *HLA-B*57:01* have an estimated 100-fold greater risk for DILI with flucloxacillin therapy. Despite this strong association, only 1 in 500–1000 patients with the *HLA-B*57:01* genotype who are undergoing flucloxacillin therapy will develop DILI. The sensitivity and specificity for pretreatment screening for *HLA-B*57:01* was high—87% and 94% respectively—but because of low incidence of flucloxacillin DILI, the PPV for the test was only 0.12%. To avoid one case of flucloxacillin-induced DILI, more than 13,500 patients would need to be screened. Furthermore, because of the low PPV, many patients who would not have developed DILI with flucloxacillin would also be denied this effective treatment [286]. At present, isolated *HLA-B*57:01* genotype screening prior to therapy is neither clinically effective nor cost-effective for routine clinical practice, but it may be of use diagnostically in patients who have developed liver manifestations where there are competing etiologies. It is interesting that *HLA-B*57:01* is also associated with abacavir hypersensitivity syndrome despite its clinical presentation being markedly different—DISI versus DILI (see above).

20.4.4.3 Ximelagatran

Ximelagatran was the first commercially available oral direct thrombin inhibitor. It was demonstrated to be equally effective as the conventional anticoagulant (warfarin) without the requirement of regular monitoring. Routine clinical and preclinical toxicological studies did not identify ximelagatran as a cause of DILI, but the drug was withdrawn from clinical development in 2007 because long-term analyses reported an association with hepatotoxicity in some patients [287]. Following drug withdrawal, a GWAS identified an association in both the exploratory and replication cohorts between ximelagatran-induced DILI and HLA alleles *DRB1*07:01* (OR = 4.4, $p_{min} = 9.1 \times 10^{-6}$, 95% CI 2.2–8.9) and *DQA1*02:01* (OR = 4.4, $p_{min} = 1.3 \times 10^{-5}$, 95% CI 2.2–8.8) [273].

It is interesting that the frequency of *HLA-DRB1*07:01* has a carrier frequency of >10% in European populations compared with 0.3% in Japanese populations [288]. The allele frequency distribution is consistent with the observation that ximelagatran DILI has a higher incidence in European populations compared with Asian populations and represents a situation similar to that of CBZ SJS/TEN and *HLA-B*15:02* [289]. The association between ximelagatran DILI and HLA alleles is surprising because the clinical phenotype of injury does not suggest an immune component. Results from lymphocyte transformation and major histocompatibility binding (MHC) tests suggest that ximelagatran is immunogenic and able to inhibit the binding of peptides to *HLA-DRB1*0701* [273]. A pharmacogenetic screening test for *HLA-DRB1*07:01* before ximelagtran therapy would have 47–50% sensitivity, 83–97% specificity, 19–59% PPV, and 95–96% NPV. Assuming an incidence of elevated alanine aminotransferases (>3 times the upper limit of normal) in 8% of patients prescribed long-term treatment with ximelagatran, between 20–24 patients would need to be screened to prevent one case of ximelagatran DILI [286,289]. Such numbers compare favorably to other screening tests, such as mammography in the identification of breast cancer [290]. These results are only based on one study, which cannot be replicated because the drug has been withdrawn, and so they need to be interpreted with caution.

20.4.4.4 Lumiracoxib

Lumiracoxib, a cyclo-oxygenase-2 selective inhibitor for treatment of pain and osteoarthritis, was withdrawn from the market or not approved following cases of severe hepatotoxicity in 2007. A GWAS compared lumiracoxib-treated patients with DILI against lumiraxocib-tolerant patients and reported a significant association with the MHC class II haplotype: *HLA-DRB1*15:01-HLA-DQB1*06:02-HLA-DRB5*01:01-HLA-DQA1*01:02* (OR = 5.0, $p = 6.8 \times 10^{-25}$, 95%

CI 3.6–7.0) [274]. Interestingly, the allele *HLA-DRB1*15:01* has also been implicated in AC-induced DILI, which suggests a possible overlap in pathogenesis of DILI despite differences in drug structure and DILI's clinical phenotype.

20.4.4.5 Other Drugs

Candidate gene studies have identified HLA associations with nevirapine-, lapatinib-, and ticlopidine-induced DILI [291–295]. Nevirapine is a widely prescribed non-nucleoside reverse transcriptase inhibitor for the treatment of HIV infection; it is associated with significantly increased frequency of hepatotoxicity compared with other antiretroviral therapies [296]. *HLA-DRB1*01:01* was significantly associated with nevirapine-induced DILI in a White Australian cohort with high CD4 T-cell counts ($OR = 3.02$, $p_c = 0.0074$, 95% CI 1.66–5.49). The association did not reach significance in Asian or black populations because of a low background frequency of the *HLA-DRB1*01:01* genotype and smaller sample sizes [292].

Lapatinib, a tyrosine kinase inhibitor, is used in the treatment of patients with advanced and metastatic breast cancer; it is associated with hepatobiliary disorder in up to 8% of patients [297]. *HLA-DQA1*02:01* was more common among patients with elevated alanine transaminase on lapatinib therapy compared with lapatinib-treated controls ($OR = 2.6$, $p = 0.03$, 95% CI 1.1–5.7). The association was confirmed in a replication cohort ($OR = 9.0$, $p = 0.0008$, 95% CI 3.2–27.4). As a predictor of lapatinib hepatotoxicity, *HLA-DQA1*02:01* had NPV and PPV of 0.97 and 0.17, respectively [293].

Ticlopidine is an antiplatelet agent that is associated with DILI; it is primarily of a cholestatic phenotype, with increased frequency in Japanese patients relative to Caucasian patients [298]. An initial case study in 22 Japanese patients with ticlopidine-induced DILI reported a significant association with *HLA-A*33:03* ($OR = 13.0$, $p_c = 1.24 \times 10^{-5}$, 95% CI 4.4–38.6) [294]. A second study reported that Japanese patients were most susceptible to ticlopidine DILI if they were positive for both *HLA-A*33:03* and *CYP2B6*1H* or *1J* ($OR = 38.8$, $p < 0.001$, 95% CI 8.8–196.0) compared with patients possessing only *HLA-A*33:03* ($OR = 15.5$, $p = < 0.001$, 95% CI 4.9–59.3) [295]. CYP2B6 is one of the key enzymes responsible for the formation of the toxic S-oxide metabolite of ticlopidine, which is hypothesized to be responsible for liver injury [299]. This study is the first to report a significant association between susceptibility to ADRs and a combination of polymorphisms in drug metabolizing enzymes and HLA.

20.5 CONCLUSION

Significant developments and breakthroughs during the past decade have increased our understanding and knowledge of genetic susceptibility to ADRs. However, apart from isolated examples such as TPMT genotyping and *HLA-B*57:01* screening in abacavir therapy, no pharmacogenetic test for ADR susceptibility has been translated into clinical practice. Clinical implementation of pharmacogenetic biomarker tests is difficult but can be split into four phases: T1 (discovery); T2 (clinical validity and utility); T3 (implementation); and T4 (effect on public health) [300]. Much of the research discussed in this chapter remains trapped in T1. This is for a multitude of reasons, including poor study design, inadequate sample sizes, lack of replication, and heterogeneity among patient populations.

Despite the many genetic associations being reported, replication has proven difficult because of the large numbers of patients and controls required. Many studies are retrospective, where patient phenotypes are recorded inaccurately and other crucial factors (e.g., drug dosage/timing, concomitant medications) are missing. Reports have estimated that only 30–50% of reported cases of DILI are correctly diagnosed or classified [301,302]. In future studies, phenotype standardization and development of diagnostic algorithms will help to improve diagnosis [193,266]. The arrival of international networks and consortia—the Drug-Induced Liver Injury Network in the U.S. (DILIN; https://dilin.dcri.duke.edu), the European collaboration to establish a case-control DNA collection for studying the genetic basis of adverse drug reactions, known as EUDRAGENE (www.eudragene.org), and the International Serious Adverse Events Consortium (SAEC; www.saeconsortium.org), among others—will enable larger samples to be collected across multiple populations and will ensure more accurate data recording.

Accurate phenotype recording is crucial because different phenotypes are likely to have diverse pathophysiological mechanisms, and combining diverse phenotypes reduces the power of a study to detect significant associations. This has been highlighted by the phenotype specificity of *HLA-B*15:02*, which is significantly associated with CBZ-induced SJS/TEN but not with HSS or MPE [221]. As well as being phenotype specific, *HLA-B*15:02* and CBZ-SJS/TEN is ethnicity specific, with the association relevant in Asian populations but not in Europeans, Japanese, or Koreans [230]. Similarly, the association between AC-induced DILI and *HLA-DRB1*15:*01 has been reported in several European populations but not in Spanish patients [275]. Susceptibility to ticlopidine hepatotoxicity is associated with an increased frequency of *HLA-A*33:03* in Japanese, but not Caucasian, populations [295]. Ethnic specificity is believed to be due to different background frequencies of the risk HLA alleles. The frequency of *HLA-A*33:03* in Japanese patients (0.1280) is significantly higher than in English (0.003), Portuguese (0.011), or Italian (0.000) populations (www.allelefrequencies.net).

In some cases, combinations of genetic polymorphisms may increase susceptibility to ADRs, as highlighted by the increased risk of ticlopidine-DILI in Japanese patients positive for both *HLA-A*33:03* and *CYP2B6*1H* or **1J*, compared with carriage of the HLA allele alone [295]. It is likely that most ADRs are polygenic and that each gene has a small overall effect. By recruiting large numbers of patients and screening for multiple genetic polymorphisms (metabolizing enzymes, drug transporters, drug receptors, immune system variation), we may be able to improve accuracy of ADR prediction in the future. A pharmacogenetic score combining three SNPs accurately predicted and stratified those patients with stable coronary artery disease into those most likely to benefit or, conversely, suffer harm from perindopril therapy, highlighting the utility of such an approach [303].

At present, next-generation sequencing (NGS) has not been utilized to investigate the role of genetics in ADRs. NGS allows analysis of rare variants, which cannot be detected by GWAS, that have been postulated to have larger effect sizes and more likely to be functionally relevant compared with common variants [304]. Epigenetics is the study of mitotically heritable changes in gene expression that are not attributable to nucleic acid sequence alterations. A recent review article discussed the role of epigenetics in ADRs, concluding that, with the advent of NGS, we are likely to significantly increase our understanding of the role of the epigenome in drug response [305,306].

Many proposed pharmacogenetic tests for ADR susceptibility have high NPVs and low PPVs because of the low incidence of some ADRs [211,230,286]. This means that a significant number of patients who test positive for the suspected allele will never develop hypersensitivity, or that very large numbers of patients need to be tested in order to avoid one ADR diagnosis. This has major clinical implications because patients may be denied optimal treatment or require more costly alternative therapies, or because the pharmacogenetic test itself is not cost-effective [286]. Prospective randomized trials to assess the utility of preprescription genetic testing will be difficult to undertake in cases where the incidence of ADR is very low. This raises the question as to the type of evidence needed to demonstrate a test's clinical utility that will be acceptable to prescribing clinicians. However, for other phenotypes that require dose adjustment or where the phenotype of interest is common, it may be possible to undertake randomized trials, as is currently being done with warfarin.

Any genetic screening test for ADRs must also demonstrate cost-effectiveness in clinical practice. ADRs that have severe clinical or economic consequences and have a clear association between genotype and phenotype, and for which there is a rapid and inexpensive genetic test with a high frequency of associated variant gene—these are more likely to be cost-effective [307].

Vanderbilt University has been developing a very large biobank linking phenotype data from clinical records to genotype data gathered from routine patient blood samples (which otherwise would have been discarded) [308]. The biobank enables rapid collection of samples, 700–900 per week, diverse phenotype recruitment, and active real-time monitoring of the effects of changes in drug choice or dosage and the chosen drug's effect on clinical outcome. In Canada, a comprehensive nationwide surveillance network—the Canadian Pharmacogenomics Network for Drug Safety (CPNDS)—was established with the goal of identifying patients with ADRs and ensuring collection of relevant biological samples and accurate clinical histories [309]. It successfully identified novel biomarkers for cisplatin-induced deafness and codeine-induced infant mortality.

Moving forward, active ADR surveillance is necessary to increase identification and recruitment of patients suffering from ADRs. In order to translate pharmacogenetic associations into successful clinical practice, collaboration will be required between scientists, clinicians, government agencies, and pharmaceutical and technology companies.

ABBREVIATIONS AND ACRONYMS

6-MP	6-mercaptopurine
AC	amoxicillin-clavulanate
ADR	adverse drug reaction
AGEP	acute generalized exanthematous pustulosis
AHS	abacavir hypersensitivity syndrome
ALL	acute lymphoblastic leukemia
AZA	azathioprine
CBZ	carbamazepine
CHCT	caffeine halothane contracture test
CK	creatine kinase
DILI	drug-induced liver injury
DISI	drug-induced skin injury
DZ	dizygotic
EM	extensive metabolizer
EMHG	European Malignant Hyperthermia Group
eQTL	expression quantitative loci
FDA	U.S. Food and Drug Administration
GWAS	genome-wide association study
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
HSS	hypersensitivity syndrome
IM	intermediate metabolizer
INR	international normalized ratio
IVCT	*in vitro* contracture test
IWPC	International Warfarin Pharmacogenetics Consortium
LTG	lamotrigine
MH	malignant hyperthermia
MHC	major histocompatibility complex
MPE	maculopapular exanthema
MZ	monozygotic
NAMHG	North American Malignant Hyperthermia Group
NAT	N-acetyltransferase
NGS	next-generation sequencing
NHS	National Health Service
OXC	oxcarbazepine

PM	poor metabolizer
RCT	randomized control trial
SJS	Stevens-Johnson syndrome
SMX	sulfamethoxazole
TB	tuberculosis
TEN	toxic epidermal necrolysis
TPMT	thiopurine methyl S-transferase
UGT	uridine diphosphate glucuronyltransferase
URM	ultrarapid metabolizer
VKOR	vitamin K epoxide reductase
VKORC	vitamin k epoxide reductase subunit 1

REFERENCES

[1] Aronson JK, Ferner RE. Clarification of terminology in drug safety. Drug Saf 2005;28(10):851–70.

[2] Rawlins MD. Clinical pharmacology. Adverse reactions to drugs. Br Med J (Clin Res Ed) 1981;282(6268):974–6.

[3] MHRA. Adverse drug reactions. February 2013. Available from: http://www.mhra.gov.uk/Safetyinformation/Howwemonitorthesafetyofproducts/Medicines/TheYellowCardScheme/Informationforhealthcareprofessionals/Adversedrugreactions/index.htm [cited 6/09/2013].

[4] Aronson JK, Ferner RE. Joining the DoTS: new approach to classifying adverse drug reactions. BMJ 2003;327:1222.

[5] Friedmann PS, Moss C, Shuster S, Simpson JM. Quantitative relationships between sensitizing dose of DNCB and reactivity in normal subjects. Clin Exp Immunol 1983;53(3):709–15.

[6] Kelkar PS, Li JTC. Cephalosporin allergy. N Engl J Med 2001;345(11):804–9.

[7] Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 2004;7456:15.

[8] Davies EC, Green CF, Taylor S, Williamson PR, Mottram DR, Pirmohamed M. Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. PLoS ONE 2009;4(2):e4439.

[9] Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;279(15):1200–5.

[10] Wester K, Jonsson AK, Spigset O, Druid H, Hagg S. Incidence of fatal adverse drug reactions: a population based study. Br J Clin Pharmacol 2008;65(4):573–9.

[11] Smyth RM, Gargon E, Kirkham J, Cresswell L, Golder S, Smyth R, et al. Adverse drug reactions in children–a systematic review. PLoS ONE 2012;7(3):e24061.

[12] Gallagher RM, Mason JR, Bird KA, Kirkham JJ, Peak M, Williamson PR, et al. Adverse drug reactions causing admission to a paediatric hospital. PLoS ONE 2012;7(12):e50127.

[13] Evans DA. Genetic variations in the acetylation of isoniazid and other drugs. Ann NY Acad Sci 1968;151(2):723–33.

[14] Vatsis KP, Martell KJ, Weber WW. Diverse point mutations in the human gene for polymorphic N-acetyltransferase. Proc Natl Acad Sci USA 1991;88(14):6333–7.

[15] Upton AM, Mushtaq A, Victor TC, Sampson SL, Sandy J, Smith DM, et al. Arylamine N-acetyltransferase of mycobacterium tuberculosis is a polymorphic enzyme and a site of isoniazid metabolism. Mol Microbiol 2001;42(2):309–17.

[16] Gu J, Liang D, Wang Y, Lu C, Wu X. Effects of N-acetyl transferase 1 and 2 polymorphisms on bladder cancer risk in Caucasians. Mutat Res 2005;581(1–2):97–104.

[17] Boukouvala S, Fakis G. Arylamine N-acetyltransferases: what we learn from genes and genomes. Drug Metab Rev 2005;37(3):511–64.

[18] Vessell E, Penno M. A new polymorphism of hepatic drug oxidation in man: family studies on rates of formation of antipyrine metabolites. Biochem Soc Trans 1984;12:74–8.

[19] Nebert DW, Zhang G, Vesell ES. Genetic risk prediction: individualized variability in susceptibility to toxicants. Annu Rev Pharmacol Toxicol 2013;53(1):355–75.

[20] Ingelman-Sundberg M, Oscarson M, McLellan RA. Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. Trends Pharmacol Sci 1999;20(8):342–9.

[21] Meyer UA, Zanger UM. Molecular mechanisms of genetic polymorphisms of drug metabolism. Annu Rev Pharmacol Toxicol 1997;37:269–96.

[22] Pirmohamed M, Park BK. Genetic susceptibility to adverse drug reactions. Trends Pharmacol Sci 2001;22(6):298–305.

[23] CYPalleles. Home page of the human cytochrome P450 allele nomenclature committee. Available from: http://www.cypalleles.ki.se/ [cited 13/08/2013].

[24] Borg KO, Carlsson E, Hoffmann KJ, Jonsson TE, Thorin H, Wallin B. Metabolism of metoprolol-(3-h) in man, the dog and the rat. Acta Pharmacol Toxicol 1975;36(Suppl. 5):125–35.

[25] McGourty JC, Silas JH, Lennard MS, Tucker GT, Woods HF. Metoprolol metabolism and debrisoquine oxidation polymorphism–population and family studies. Br J Clin Pharmacol 1985;20(6):555–66.

[26] Lennard MS, Silas JH, Freestone S, Ramsay LE, Tucker GT, Woods HF. Oxidation phenotype–a major determinant of metoprolol metabolism and response. N Engl J Med 1982;307(25):1558–60.

[27] Rau T, Heide R, Bergmann K, Wuttke H, Werner U, Feifel N, et al. Effect of the CYP2D6 genotype on metoprolol metabolism persists during long-term treatment. Pharmacogenetics 2002;12(6):465–72.

[28] Wuttke H, Rau T, Heide R, Bergmann K, Bohm M, Weil J, et al. Increased frequency of cytochrome P450 2D6 poor metabolizers among patients with metoprolol-associated adverse effects. Clin Pharmacol Ther 2002;72(4):429–37.

[29] Fux R, Morike K, Prohmer AM, Delabar U, Schwab M, Schaeffeler E, et al. Impact of CYP2D6 genotype on adverse effects during treatment with metoprolol: a prospective clinical study. Clin Pharmacol Ther 2005;78(4):378–87.

[30] Baudhuin LM, Miller WL, Train L, Bryant S, Hartman KA, Phelps M, et al. Relation of ADRB1, CYP2D6, and UGT1A1 polymorphisms with dose of, and response to, carvedilol or metoprolol therapy in patients with chronic heart failure. Am J Cardiol 2010;106(3):402–8.

[31] Lennard MS, Tucker GT, Silas JH, Woods HF. Debrisoquine polymorphism and the metabolism and action of metoprolol, timolol, propranolol and atenolol. Xenobiotica 1986;16(5):435–47.

[32] Giessmann T, Modess C, Hecker U, Zschiesche M, Dazert P, Kunert-Keil C, et al. CYP2D6 genotype and induction of intestinal drug transporters by rifampin predict presystemic clearance of carvedilol in healthy subjects. Clin Pharmacol Ther 2004;75(3):213–22.

[33] Kirchheiner J, Schmidt H, Tzvetkov M, Keulen JT, Lotsch J, Roots I, et al. Pharmacokinetics of codeine and its metabolite morphine in ultra-rapid metabolizers due to CYP2D6 duplication. Pharmacogenomics J 2007;7(4):257–65.

[34] FDA. Codeine. Available from: http://www.fda.gov/Drugs/DrugSafety/ucm339112.htm [cited 13/08/2013].

[35] Ciszkowski C, Madadi P, Phillips MS, Lauwers AE, Koren G. Codeine, ultrarapid-metabolism genotype, and postoperative death. N Engl J Med 2009;361(8):827–8.

[36] Kelly LE, Rieder M, van den Anker J, Malkin B, Ross C, Neely MN, et al. More codeine fatalities after tonsillectomy in North American children. Pediatrics 2012;129(5):e1343–7.

[37] MHRA. Safety update codeine. Available from: http://www.mhra.gov.uk/Safetyinformation/DrugSafetyUpdate/CON296400) [cited 13/08/2013].

[38] Koren G, Cairns J, Chitayat D, Gaedigk A, Leeder SJ. Pharmacogenetics of morphine poisoning in a breastfed neonate of a codeine-prescribed mother. Lancet 2006;368(9536):704.

[39] Meny RG, Naumburg EG, Alger LS, Brill-Miller JL, Brown S. Codeine and the breastfed neonate. J Hum Lact 1993;9(4):237–40.

[40] Chaudhry AS, Urban TJ, Lamba JK, Birnbaum AK, Remmel RP, Subramanian M, et al. CYP2C9*1B promoter polymorphisms, in linkage with CYP2C19*2, affect phenytoin autoinduction of clearance and maintenance dose. J Pharmacol Exp Ther 2010;332(2):599–611.

[41] Kupfer A, Preisig R. Pharmacogenetics of mephenytoin: a new drug hydroxylation polymorphism in man. Eur J Clin Pharmacol 1984;26(6):753–9.

[42] Aynacioglu AS, Brockmoller J, Bauer S, Sachse C, Guzelbey P, Ongen Z, et al. Frequency of cytochrome P450 CYP2C9 variants in a Turkish population and functional relevance for phenytoin. Br J Clin Pharmacol 1999;48(3):409–15.

[43] Gotoh O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J Biol Chem 1992;267(1):83–90.

[44] Odani A, Hashimoto Y, Otsuki Y, Uwai Y, Hattori H, Furusho K, et al. Genetic polymorphism of the CYP2C subfamily and its effect on the pharmacokinetics of phenytoin in Japanese patients with epilepsy. Clin Pharmacol Ther 1997;62(3):287–92.

[45] Mamiya K, Ieiri I, Shimamoto J, Yukawa E, Imai J, Ninomiya H, et al. The effects of genetic polymorphisms of CYP2C9 and CYP2C19 on phenytoin metabolism in Japanese adult patients with epilepsy: studies in stereoselective hydroxylation and population pharmacokinetics. Epilepsia 1998;39(12):1317–23.

[46] Depondt C, Godard P, Espel RS, Da Cruz AL, Lienard P, Pandolfo M. A candidate gene study of antiepileptic drug tolerability and efficacy identifies an association of CYP2C9 variants with phenytoin toxicity. Eur J Neurol 2011;18(9):1159–64.

[47] Tate SK, Depondt C, Sisodiya SM, Cavalleri GL, Schorge S, Soranzo N, et al. Genetic predictors of the maximum doses patients receive during clinical use of the anti-epileptic drugs carbamazepine and phenytoin. Proc Natl Acad Sci USA 2005;102(15):5507–12.

[48] Twardowschy CA, Werneck LC, Scola RH, Borgio JG, De Paola L, Silvado C. The role of CYP2C9 polymorphisms in phenytoin-related cerebellar atrophy. Seizure 2013;22(3):194–7.

[49] Cacabelos R, Hashimoto R, Takeda M. Pharmacogenomics of antipsychotics efficacy for schizophrenia. Psychiatry Clin Neurosci 2011;65(1):3–19.

[50] Fleeman N, Dundar Y, Dickson R, Jorgensen A, Pushpakom S, McLeod C, et al. Cytochrome P450 testing for prescribing antipsychotics in adults with schizophrenia: systematic review and meta-analyses. Pharmacogenomics J 2011;11(1):1–14.

[51] Chou WH, Yan FX, de Leon J, Barnhill J, Rogers T, Cronin M, et al. Extension of a pilot study: impact from the cytochrome P450 2D6 polymorphism on outcome and costs associated with severe mental illness. J Clin Psychopharmacol 2000;20(2):246–51.

[52] Basile VS, Ozdemir V, Masellis M, Walker ML, Meltzer HY, Lieberman JA, et al. A functional polymorphism of the cytochrome P450 1A2 (CYP1A2) gene: association with tardive dyskinesia in schizophrenia. Mol Psychiatry 2000;5(4):410–7.

[53] Schulze TG, Schumacher J, Muller DJ, Krauss H, Alfter D, Maroldt A, et al. Lack of association between a functional polymorphism of the cytochrome P450 1A2 (CYP1A2) gene and tardive dyskinesia in schizophrenia. Am J Med Genet 2001;105(6):498–501.

[54] Lennard L, van Loon JA, Weinshilboum RM. Pharmacogenetics of acute azathioprine toxicity: relationship to thiopurine methyltransferase genetic polymorphism. Clin Pharmacol Ther 1989;46(2):149–54.

[55] Stanulla M, Schaeffeler E, Flohr T, Cario G, Schrauder A, Zimmermann M, et al. Thiopurine methyltransferase (TPMT) genotype and early treatment response to mercaptopurine in childhood acute lymphoblastic leukemia. JAMA 2005;293(12):1485–9.

[56] Hindorf U, Lindqvist M, Peterson C, Soderkvist P, Strom M, Hjortswang H, et al. Pharmacogenetics during standardised initiation of thiopurine treatment in inflammatory bowel disease. Gut 2006;55(10):1423–31.

[57] Luce JK, Frenkel, E.P., Vietti TJ, et al. Clinical studies of 6-methylmercaptopurine riboside (NSC-40774) in acute leukaemia. Cancer Chemother Rep 1967;51: 535–46.

[58] Relling MV, Hancock ML, Rivera GK, Sandlund JT, Ribeiro RC, Krynetski EY, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. J Natl Cancer Inst 1999;91(23):2001–8.

[59] Winter JW, Gaffney D, Shapiro D, Spooner RJ, Marinaki AM, Sanderson JD, et al. Assessment of thiopurine methyltransferase enzyme activity is superior to genotype in predicting myelosuppression following azathioprine therapy in patients with inflammatory bowel disease. Aliment Pharmacol Ther 2007;25(9):1069–77.

[60] Schwab M, Schaffeler E, Marx C, Fischer C, Lang T, Behrens C, et al. Azathioprine therapy and adverse drug reactions in patients with inflammatory bowel disease: impact of thiopurine S-methyltransferase polymorphism. Pharmacogenetics 2002;12(6):429–36.

[61] Hindorf U, Appell ML. Genotyping should be considered the primary choice for pre-treatment evaluation of thiopurine methyltransferase function. J Crohns Colitis 2012;6(6):655–9.

[62] Hon YY, Fessing MY, Pui CH, Relling MV, Krynetski EY, Evans WE. Polymorphism of the thiopurine S-methyltransferase gene in African-Americans. Hum Mol Genet 1999;8(2):371–6.

[63] Krynetski EY, Schuetz JD, Galpin AJ, Pui CH, Relling MV, Evans WE. A single point mutation leading to loss of catalytic activity in human thiopurine S-methyltransferase. Proc Natl Acad Sci USA 1995;92(4):949–53.

[64] Otterness D, Szumlanski C, Lennard L, Klemetsdal B, Aarbakke J, Park-Hah JO, et al. Human thiopurine methyltransferase pharmacogenetics: gene sequence polymorphisms. Clin Pharmacol Ther 1997;62(1):60–73.

[65] Yates CR, Krynetski EY, Loennechen T, Fessing MY, Tai HL, Pui CH, et al. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. Ann Intern Med 1997;126(8):608–14.

[66] Haglund S, Lindqvist M, Almer S, Peterson C, Taipalensuu J. Pyrosequencing of TPMT alleles in a general Swedish population and in patients with inflammatory bowel disease. Clin Chem 2004;50(2):288–95.

[67] Relling MV, Gardner EE, Sandborn WJ, Schmiegelow K, Pui CH, Yee SW, et al. Clinical pharmacogenetics implementation consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing. Clin Pharmacol Ther 2011;89(3):387–91.

[68] Ansari A, Arenas M, Greenfield SM, Morris D, Lindsay J, Gilshenan K, et al. Prospective evaluation of the pharmacogenetics of azathioprine in the treatment of inflammatory bowel disease. Aliment Pharmacol Ther 2008;28(8):973–83.

[69] Donnan JR, Ungar WJ, Mathews M, Rahman P. Systematic review of thiopurine methyltransferase genotype and enzymatic testing strategies. Ther Drug Monit 2011;33(2):192–9.

[70] Colombel JF, Ferrari N, Debuysere H, Marteau P, Gendre JP, Bonaz B, et al. Genotypic analysis of thiopurine S-methyltransferase in patients with Crohn's disease and severe myelosuppression during azathioprine therapy. Gastroenterology 2000;118(6):1025–30.

[71] FDA. Azathioprine. Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/016324s034s035lbl.pdf [cited 11/02/14].

[72] Kornbluth A, Sachar DB. Practice parameters committee of the American college of G. Ulcerative colitis practice guidelines in adults: American college of gastroenterology, practice parameters committee. Am J Gastroenterol 2010;105(3):501–23. quiz 24

[73] Ford LT, Berg JD. Thiopurine S-methyltransferase (TPMT) assessment prior to starting thiopurine drug treatment; a pharmacogenomic test whose time has come. J Clin Pathol 2010;63(4):288–95.

[74] Carter MJ, Lobo AJ, Travis SP. Ibd Section BSoG. Guidelines for the management of inflammatory bowel disease in adults. Gut 2004;53(Suppl. 5):V1–V16.

[75] Winter J, Walker A, Shapiro D, Gaffney D, Spooner RJ, Mills PR. Cost-effectiveness of thiopurine methyltransferase genotype screening in patients about to commence azathioprine therapy for treatment of inflammatory bowel disease. Aliment Pharmacol Ther 2004;20(6):593–9.

[76] Oh KT, Anis AH, Bae SC. Pharmacoeconomic analysis of thiopurine methyltransferase polymorphism screening by polymerase chain reaction for treatment with azathioprine in Korea. Rheumatology 2004;43(2):156–63.

[77] Dubinsky MC, Reyes E, Ofman J, Chiou CF, Wade S, Sandborn WJ. A cost-effectiveness analysis of alternative disease management strategies in patients with Crohn's disease treated with azathioprine or 6-mercaptopurine. Am J Gastroenterol 2005;100(10):2239–47.

[78] Marra CA, Esdaile JM, Anis AH. Practical pharmacogenetics: the cost effectiveness of screening for thiopurine s-methyltransferase polymorphisms in patients with rheumatological conditions treated with azathioprine. J Rheumatol 2002;29(12):2507–12.

[79] Priest VL, Begg EJ, Gardiner SJ, Frampton CM, Gearry RB, Barclay ML, et al. Pharmacoeconomic analyses of azathioprine, methotrexate and prospective pharmacogenetic testing for the management of inflammatory bowel disease. PharmacoEconomics 2006;24(8):767–81.

[80] van den Akker-van Marle ME, Gurwitz D, Detmar SB, Enzing CM, Hopkins MM, Gutierrez de Mesa E, et al. Cost-effectiveness of pharmacogenomics in clinical practice: a case study of thiopurine methyltransferase genotyping in acute lymphoblastic leukemia in Europe. Pharmacogenomics 2006;7(5):783–92.

[81] Donnan JR, Ungar WJ, Mathews M, Hancock-Howard RL, Rahman P. A cost effectiveness analysis of thiopurine methyltransferase testing for guiding 6-mercaptopurine dosing in children with acute lymphoblastic leukemia. Pediatr Blood Cancer 2011;57(2):231–9.

[82] Ross CJ, Katzov-Eckert H, Dube MP, Brooks B, Rassekh SR, Barhdadi A, et al. Genetic variants in TPMT and COMT are associated with hearing loss in children receiving cisplatin chemotherapy. Nat Genet 2009;41(12):1345–9.

[83] Pussegoda K, Ross CJ, Visscher H, Yazdanpanah M, Brooks B, Rassekh SR, et al. Replication of TPMT and ABCC3 Genetic variants highly associated with cisplatin-induced hearing loss in children. Clin Pharmacol Ther 2013;94(2):243–51.

[84] FDA. Cisplatin. Available from: http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm287714.htm [cited 13/09/13].

[85] Walker K, Ginsberg G, Hattis D, Johns DO, Guyton KZ, Sonawane B. Genetic polymorphism in N-acetyltransferase (NAT): population distribution of NAT1 and NAT2 activity. J Toxicol Environ Health B Crit Rev 2009;12(5–6):440–72.

[86] Higuchi N, Tahara N, Yanagihara K, Fukushima K, Suyama N, Inoue Y, et al. NAT2 6A, a haplotype of the N-acetyltransferase 2 gene, is an important biomarker for risk of anti-tuberculosis drug-induced hepatotoxicity in Japanese patients with tuberculosis. World J Gastroenterol 2007;13(45):6003–8.

[87] Bailliet G, Santos MR, Alfaro EL, Dipierri JE, Demarchi DA, Carnese FR, et al. Allele and genotype frequencies of metabolic genes in Native Americans from Argentina and Paraguay. Mutat Res 2007;627(2):171–7.

[88] Yamada S, Tang M, Richardson K, Halaschek-Wiener J, Chan M, Cook VJ, et al. Genetic variations of NAT2 and CYP2E1 and isoniazid hepatotoxicity in a diverse population. Pharmacogenomics 2009;10(9):1433–45.

[89] Ohno M, Yamaguchi I, Yamamoto I, Fukuda T, Yokota S, Maekura R, et al. Slow N-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity. Int J Tuberc Lung Dis 2000;4(3):256–61.

[90] Gurumurthy P, Krishnamurthy MS, Nazareth O, Parthasarathy R, Sarma GR, Somasundaram PR, et al. Lack of relationship between hepatic toxicity and acetylator phenotype in three thousand South Indian patients during treatment with isoniazid for tuberculosis. Am Rev Respir Dis 1984;129(1):58–61.

[91] Garcia-Closas M, Hein DW, Silverman D, Malats N, Yeager M, Jacobs K, et al. A single nucleotide polymorphism tags variation in the arylamine N-acetyltransferase 2 phenotype in populations of European background. Pharmacogenet Genomics 2011;21(4):231–6.

[92] Ho HT, Wang TH, Hsiong CH, Perng WC, Wang NC, Huang TY, et al. The NAT2 tag SNP rs1495741 correlates with the susceptibility of antituberculosis drug-induced hepatotoxicity. Pharmacogenet Genomics 2013;23(4):200–7.

[93] FDA. Rifater. Available from: http://www.fda.gov/Safety/MedWatch/SafetyInformation/ucm235441.htm [cited 12/09/13].

[94] Schneider MM, Nielsen TL, Nelsing S, Hoepelman AI, Eeftinck Schattenkerk JK, van der Graaf Y, et al. Efficacy and toxicity of two doses of trimethoprim-sulfamethoxazole as primary prophylaxis against pneumocystis carinii pneumonia in patients with human immunodeficiency virus. Dutch AIDS treatment group. J Infect Dis 1995;171(6):1632–6.

[95] Bigby M, Jick S, Jick H, Arndt K. Drug-induced cutaneous reactions. A report from the Boston collaborative drug surveillance program on 15,438 consecutive inpatients, 1975 to 1982. JAMA 1986;256(24):3358–63.

[96] Jaffe HS, Abrams DI, Ammann AJ, Lewis BJ, Golden JA. Complications of co-trimoxazole in treatment of AIDS-associated pneumocystis carinii pneumonia in homosexual men. Lancet 1983;2(8359):1109–11.

[97] Rieder MJ, Shear NH, Kanee A, Tang BK, Spielberg SP. Prominence of slow acetylator phenotype among patients with sulfonamide hypersensitivity reactions. Clin Pharmacol Ther 1991;49(1):13–17.

[98] Carr A, Gross AS, Hoskins JM, Penny R, Cooper DA. Acetylation phenotype and cutaneous hypersensitivity to trimethoprim-sulfamethoxazole in HIV-infected patients. AIDS 1994;8(3):333–7.

[99] Kaufmann GR, Wenk M, Taeschner W, Peterli B, Gyr K, Meyer UA, et al. N-acetyltransferase 2 polymorphism in patients infected with human immunodeficiency virus. Clin Pharmacol Ther 1996;60(1):62–7.

[100] Pirmohamed M, Alfirevic A, Vilar J, Stalford A, Wilkins EG, Sim E, et al. Association analysis of drug metabolizing enzyme gene polymorphisms in HIV-positive patients with co-trimoxazole hypersensitivity. Pharmacogenetics 2000;10(8):705–13.

[101] Delomenie C, Grant DM, Mathelier-Fusade P, Jacomet C, Leynadier F, Jacqz-Aigrain E, et al. N-acetylation genotype and risk of severe reactions to sulphonamides in AIDS patients. Br J Clin Pharmacol 1994;38(6):581–2.

[102] Wolkenstein P, Loriot MA, Aractingi S, Cabelguenne A, Beaune P, Chosidow O. Prospective evaluation of detoxification pathways as markers of cutaneous adverse reactions to sulphonamides in AIDS. Pharmacogenetics 2000;10(9):821–8.

[103] Alfirevic A, Stalford AC, Vilar FJ, Wilkins EG, Park BK, Pirmohamed M. Slow acetylator phenotype and genotype in HIV-positive patients with sulfamethoxazole hypersensitivity. Br J Clin Pharmacol 2003;55(2):158–65.

[104] Kadakol A, Ghosh SS, Sappal BS, Sharma G, Chowdhury JR, Chowdhury NR. Genetic lesions of bilirubin uridine-diphosphoglucuronate glucuronosyltransferase (UGT1A1) causing Crigler-Najjar and Gilbert syndromes: correlation of genotype to phenotype. Hum Mutat 2000;16(4):297–306.

[105] FDA. Irinotecan. Available from: http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm236793.htm [cited 15/08/2013].

[106] Innocenti F, Undevia SD, Iyer L, Chen PX, Das S, Kocherginsky M, et al. Genetic variants in the UDP-glucuronosyltransferase 1A1 gene predict the risk of severe neutropenia of irinotecan. J Clin Oncol 2004;22(8):1382–8.

[107] Liu CY, Chen PM, Chiou TJ, Liu JH, Lin JK, Lin TC, et al. UGT1A1*28 polymorphism predicts irinotecan-induced severe toxicities without affecting treatment outcome and survival in patients with metastatic colorectal carcinoma. Cancer 2008;112(9):1932–40.

[108] Xu JM, Wang Y, Ge FJ, Lin L, Liu ZY, Sharma MR. Severe irinotecan-induced toxicity in a patient with UGT1A1*28 and UGT1A1*6 polymorphisms. WJG 2013;19(24):3899–903.

[109] Gao J, Zhou J, Li Y, Lu M, Jia R, Shen L. UGT1A1 6/28 polymorphisms could predict irinotecan-induced severe neutropenia not diarrhea in Chinese colorectal cancer patients. Med Oncol 2013;30(3):604.

[110] Levesque E, Belanger AS, Harvey M, Couture F, Jonker D, Innocenti F, et al. Refining the UGT1A haplotype associated with irinotecan-induced hematological toxicity in metastatic colorectal cancer patients treated with 5-fluorouracil/irinotecan-based regimens. J Pharmacol Exp Ther 2013;345(1):95–101.

[111] Wang Y, Shen L, Xu N, Wang JW, Jiao SC, Liu ZY, et al. UGT1A1 predicts outcome in colorectal cancer treated with irinotecan and fluorouracil. World J Gastroenterol: 2012;18(45):6635–44.

[112] Zhou CF, Ma T, Su Y, Ye ZB, Ji J, Yu YY, et al. UGT1A1 gene polymorphisms and the toxicities of FOLFIRI in Chinese Han patients with gastrointestinal cancer. Anti-Cancer Agents Med Chem 2013;13(2):235–41.

[113] Hoskins JM, Goldberg RM, Qu P, Ibrahim JG, McLeod HL. UGT1A1*28 genotype and irinotecan-induced neutropenia: dose matters. J Natl Cancer Inst 2007;99(17):1290–5.

[114] Dias MM, McKinnon RA, Sorich MJ. Impact of the UGT1A1*28 allele on response to irinotecan: a systematic review and meta-analysis. Pharmacogenomics 2012;13(8):889–99.

[115] Yousef ZR, Tandy SC, Tudor V, Jishi F, Trent RJ, Watson DK, et al. Warfarin for non-rheumatic atrial fibrillation: five year experience in a district general hospital. Heart 2004;90(11):1259–62.

[116] Johnson CE, Lim WK, Workman BS. People aged over 75 in atrial fibrillation on warfarin: the rate of major hemorrhage and stroke in more than 500 patient-years of follow-up. J Am Geriatr Soc 2005;53(4):655–9.

[117] Eble JN, West BD, Link KP. A comparison of the isomers of warfarin. Biochem Pharmacol 1966;15(7):1003–6.

[118] Gebauer M. Synthesis and structure-activity relationships of novel warfarin derivatives. Bioorg Med Chem 2007;15(6):2414–20.

[119] Fasco MJ, Principe LM. R- and S-warfarin inhibition of vitamin K and vitamin K 2,3-epoxide reductase activities in the rat. J Biol Chem 1982;257(9):4894–901.

[120] Park BK. Warfarin: metabolism and mode of action. Biochem Pharmacol 1988;37(1):19–27.

[121] Crespi CL, Miller VP. The R144C change in the CYP2C9*2 allele alters interaction of the cytochrome P450 with NADPH: cytochrome P450 oxidoreductase. Pharmacogenetics 1997;7(3):203–10.

[122] Takanashi K, Tainaka H, Kobayashi K, Yasumori T, Hosakawa M, Chiba K. CYP2C9 Ile359 and Leu359 variants: enzyme kinetic study with seven substrates. Pharmacogenetics 2000;10(2):95–104.

[123] Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, McLeod HL, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N Engl J Med 2005;352(22):2285–93.

[124] Schwarz UI, Ritchie MD, Bradford Y, Li C, Dudek SM, Frye-Anderson A, et al. Genetic determinants of response to warfarin during initial anticoagulation. N Engl J Med 2008;358(10):999–1008.

[125] Aithal GP, Day CP, Kesteven PJ, Daly AK. Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet 1999;353(9154):717–9.

[126] Higashi MK, Veenstra DL, Kondo LM, Wittkowsky AK, Srinouanprachanh SL, Farin FM, et al. Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. JAMA 2002;287(13):1690–8.

[127] D'Andrea G, D'Ambrosio RL, Di Perna P, Chetta M, Santacroce R, Brancaccio V, et al. A polymorphism in the VKORC1 gene is associated with an interindividual variability in the dose-anticoagulant effect of warfarin. Blood 2005;105(2):645–9.

[128] Hillman MA, Wilke RA, Caldwell MD, Berg RL, Glurich I, Burmester JK. Relative impact of covariates in prescribing

warfarin according to CYP2C9 genotype. Pharmacogenetics 2004;14(8):539–47.

[129] Wadelius M, Chen LY, Downes K, Ghori J, Hunt S, Eriksson N, et al. Common VKORC1 and GGCX polymorphisms associated with warfarin dose. Pharmacogenomics J 2005;5(4):262–70.

[130] Wadelius M, Sorlin K, Wallerman O, Karlsson J, Yue QY, Magnusson PK, et al. Warfarin sensitivity related to CYP2C9, CYP3A5, ABCB1 (MDR1) and other factors. Pharmacogenomics J 2004;4(1):40–8.

[131] Carlquist JF, Horne BD, Muhlestein JB, Lappe DL, Whiting BM, Kolek MJ, et al. Genotypes of the cytochrome p450 isoform, CYP2C9, and the vitamin K epoxide reductase complex subunit 1 conjointly determine stable warfarin dose: a prospective study. J Thromb Thrombolysis 2006;22(3):191–7.

[132] Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med 2009;360(8):753–64.

[133] Gage BF, Eby C, Johnson JA, Deych E, Rieder MJ, Ridker PM, et al. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. Clin Pharmacol Ther 2008;84(3):326–31.

[134] Johnson JA, Gong L, Whirl-Carrillo M, Gage BF, Scott SA, Stein CM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. Clin Pharmacol Ther 2011;90(4):625–9.

[135] Garcia DA, Lopes RD, Hylek EM. New-onset atrial fibrillation and warfarin initiation: high risk periods and implications for new antithrombotic drugs. Thromb Haemost 2010;104(6):1099–105.

[136] Michaud V, Vanier MC, Brouillette D, Roy D, Verret L, Noel N, et al. Combination of phenotype assessments and CYP2C9-VKORC1 polymorphisms in the determination of warfarin dose requirements in heavily medicated patients. Clin Pharmacol Ther 2008;83(5):740–8.

[137] Wadelius M, Chen LY, Lindh JD, Eriksson N, Ghori MJ, Bumpstead S, et al. The largest prospective warfarin-treated cohort supports genetic forecasting. Blood 2009;113(4):784–92.

[138] Finkelman BS, Gage BF, Johnson JA, Brensinger CM, Kimmel SE. Genetic warfarin dosing: tables versus algorithms. J Am Coll Cardiol 2011;57(5):612–8.

[139] Huang SW, Chen HS, Wang XQ, Huang L, Xu DL, Hu XJ, et al. Validation of VKORC1 and CYP2C9 genotypes on interindividual warfarin maintenance dose: a prospective study in Chinese patients. Pharmacogenet Genomics 2009;19(3):226–34.

[140] Caraco Y, Blotnick S, Muszkat M. CYP2C9 genotype-guided warfarin prescribing enhances the efficacy and safety of anticoagulation: a prospective randomized controlled study. Clin Pharmacol Ther 2008;83(3):460–70.

[141] Anderson JL, Horne BD, Stevens SM, Grove AS, Barton S, Nicholas ZP, et al. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. Circulation 2007;116(22):2563–70.

[142] Hillman MA, Wilke RA, Yale SH, Vidaillet HJ, Caldwell MD, Glurich I, et al. A prospective, randomized pilot trial of model-based warfarin dose initiation using CYP2C9 genotype and clinical data. Clin Med Res 2005;3(3):137–45.

[143] Sanderson S, Emery J, Higgins J. CYP2C9 gene variants, drug dose, and bleeding risk in warfarin-treated patients: a HuGEnet systematic review and meta-analysis. Genet Med 2005;7(2):97–104.

[144] Limdi NA, McGwin G, Goldstein JA, Beasley TM, Arnett DK, Adler BK, et al. Influence of CYP2C9 and VKORC1 1173C/T genotype on the risk of hemorrhagic complications in African-American and European-American patients on warfarin. Clin Pharmacol Ther 2008;83(2):312–21.

[145] Takeuchi F, McGinnis R, Bourgeois S, Barnes C, Eriksson N, Soranzo N, et al. A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. PLoS Genet 2009;5(3):e1000433.

[146] Meckley LM, Wittkowsky AK, Rieder MJ, Rettie AE, Veenstra DL. An analysis of the relative effects of VKORC1 and CYP2C9 variants on anticoagulation related outcomes in warfarin-treated patients. Thromb Haemost 2008;100(2):229–39.

[147] Limdi NA, Arnett DK, Goldstein JA, Beasley TM, McGwin G, Adler BK, et al. Influence of CYP2C9 and VKORC1 on warfarin dose, anticoagulation attainment and maintenance among European-Americans and African-Americans. Pharmacogenomics 2008;9(5):511–26.

[148] Reitsma PH, van der Heijden JF, Groot AP, Rosendaal FR, Buller HR. A C1173T dimorphism in the VKORC1 gene determines coumarin sensitivity and bleeding risk. PLoS Med 2005; 2(10):e312.

[149] Lambert CA, Tishkoff SA. Genetic structure in African populations: implications for human demographic history. Cold Spring Harb Symp Quant Biol 2009;74:395–402.

[150] Perera MA, Cavallari LH, Limdi NA, Gamazon ER, Konkashbaev A, Daneshjou R, et al. Genetic variants associated with warfarin dose in African-American individuals: a genome-wide association study. Lancet 2013;382(9894):790–6.

[151] Eltit JM, Bannister RA, Moua O, Altamirano F, Hopkins PM, Pessah IN, et al. Malignant hyperthermia susceptibility arising from altered resting coupling between the skeletal muscle L-type Ca2+ channel and the type 1 ryanodine receptor. Proc Nat Acad Sci 2012;109(20):7923–8.

[152] Denborough MA, Lovell RRH. Anæsthetic deaths in a family. Lancet 1960;276(7140):45.

[153] Brady JE, Sun LS, Rosenberg H, Li G. Prevalence of malignant hyperthermia due to anesthesia in New York state, 2001–2005. Anesth Analg 2009;109(4):1162–6.

[154] Larach MG, Gronert GA, Allen GC, Brandom BW, Lehman EB. Clinical presentation, treatment, and complications of malignant hyperthermia in North America from 1987 to 2006. Anesth Analg 2010;110(2):498–507.

[155] Riazi S, Larach MG, Hu C, Wijeysundera D, Massey C, Kraeva N. Malignant hyperthermia in Canada: characteristics of index anesthetics in 129 malignant hyperthermia susceptible probands. Anesth Analg 2013.

[156] Krause T, Gerbershagen MU, Fiege M, Weißhorn R, Wappler F. Dantrolene—a review of its pharmacology, therapeutic use and new developments. Anaesthesia 2004;59(4):364–73.

[157] Yang T, Allen PD, Pessah IN, Lopez JR. Enhanced excitation-coupled calcium entry in myotubes is associated with expression of RyR1 malignant hyperthermia mutations. J Biol Chem 2007;282(52):37471–8.

[158] Litman RS, Rosenberg H. Malignant hyperthermia: update on susceptibility testing. JAMA 2005;293(23):2918–24.

[159] The European malignant hyperpyrexia G. A protocol for the investigation of malignant hyperpyrexia (mh) susceptibility. Br J Anaesth 1984;56(11):1267–71.

[160] Larach MG. For the North American malignant hyperthermia G. Standardization of the caffeine halothane muscle contracture test. Anesth Analg 1989;69(4):511–5.

[161] Oerding H, Brancadoro V, Cozzolino S, Ellis FR, Glauber V, Gonano EF, et al. *In vitro* contracture test for diagnosis of malignant hyperthermia following the protocol of the European MH group: results of testing patients surviving fulminant MH and unrelated low-risk subjects. Acta Anaesthesiol Scand 1997;41(8):955–66.

[162] Allen GC, Larach MG, Kunselman AR. The sensitivity and specificity of the caffeine-halothane contracture test: a report from the North American malignant hyperthermia registry. Anesthesiology 1998;88(3):579–88.

[163] Hirshey Dirksen SJ, Larach MG, Rosenberg H, Brandom BW, Parness J, Lang RS, et al. Future directions in malignant hyperthermia research and patient care. Anesth Analg 2011;113(5):1108–19.

[164] Rosenberg H, Sambuughin N, Riazi S, Dirksen R. Malignant hyperthermia susceptibility. In: Pagon R, Adam M, Bird T, editors. Gene reviews. Seattle, WA: University of Washington; 2013.

[165] Brandom BW. Genetics of malignant hyperthermia. Sci World J 2006;6:1722–30.

[166] Levano S, Vukcevic M, Singer M, Matter A, Treves S, Urwyler A, et al. Increasing the number of diagnostic mutations in malignant hyperthermia. Hum Mutat 2009;30(4):590–8.

[167] Stewart S, Hogan K, Rosenberg H, Fletcher J. Identification of the Arg1086His mutation in the alpha subunit of the voltage-dependent calcium channel (CACNA1S) in a North American family with malignant hyperthermia. Clin Genet 2001;59(3):178–84.

[168] Robinson RL, Hopkins PM. Editorial III: a breakthrough in the genetic diagnosis of malignant hyperthermia. Br J Anaesth 2001;86(2):166–8.

[169] Pietri-Rouxel F, Gentil C, Vassilopoulos S, Baas D, Mouisel E, Ferry A, et al. DHPR [alpha]1S subunit controls skeletal muscle mass and morphogenesis. EMBO J 2010;29(3):643–54.

[170] Sambuughin N, Capacchione J, Blokhin A, Bayarsaikhan M, Bina S, Muldoon S. The ryanodine receptor type 1 gene variants in African American men with exertional rhabdomyolysis and malignant hyperthermia susceptibility. Clin Genet 2009;76(6):564–8.

[171] Robinson R, Carpenter D, Shaw M-A, Halsall J, Hopkins P. Mutations in RYR1 in malignant hyperthermia and central core disease. Hum Mutat 2006;27(10):977–89.

[172] Sei Y, Sambuughin NN, Davis EJ, Sachs D, Cuenca PB, Brandom BW, et al. Malignant hyperthermia in North America: genetic screening of the three hot spots in the type I ryanodine receptor gene. Anesthesiology 2004;101(4):824–30.

[173] Urwyler A, Deufel T, McCarthy T, West S. Guidelines for molecular genetic detection of susceptibility to malignant hyperthermia. Br J Anaesth 2001;86(2):283–7.

[174] Carpenter D, Robinson RL, Quinnell RJ, Ringrose C, Hogg M, Casson F, et al. Genetic variation in RYR1 and malignant hyperthermia phenotypes. Br J Anaesth 2009;103(4):538–48.

[175] Newmark JL, Voelkel M, Brandom BW, Wu J. Delayed onset of malignant hyperthermia without creatine kinase elevation in a geriatric, ryanodine receptor type 1 gene compound heterozygous patient. Anesthesiology 2007;107(2):350–3. http://dx.doi.org/10.1097/01.anes.0000271866.37280.85.

[176] Robinson RL, Carpenter D, Halsall PJ, Iles DE, Booms P, Steele D, et al. Epigenetic allele silencing and variable penetrance of malignant hyperthermia susceptibility. Br J Anaesth 2009;103(2):220–5.

[177] McKenney JM, Davidson MH, Jacobson TA, Guyton JR. National lipid association statin safety assessment task F. final conclusions and recommendations of the national lipid association statin safety assessment task force. Am J Cardiol 2006;97(8A):89C–94C.

[178] Law M, Rudnicka AR. Statin safety: a systematic review. Am J Cardiol 2006;97(8A):52C–60C.

[179] Daly AK. Genome-wide association studies in pharmacogenomics. Nat Rev Genet 2010;11(4):241–6.

[180] Thompson PD, Clarkson PM, Rosenson RS. National lipid association statin safety task force muscle safety expert P. An assessment of statin safety by muscle experts. Am J Cardiol 2006;97(8A):69C–76C.

[181] McClure DL, Valuck RJ, Glanz M, Murphy JR, Hokanson JE. Statin and statin-fibrate use was significantly associated with increased myositis risk in a managed care population. J Clin Epidemiol 2007;60(8):812–8.

[182] Pasanen MK, Fredrikson H, Neuvonen PJ, Niemi M. Different effects of SLCO1B1 polymorphism on the pharmacokinetics of atorvastatin and rosuvastatin. Clin Pharmacol Ther 2007;82(6):726–33.

[183] Ho RH, Choi L, Lee W, Mayo G, Schwarz UI, Tirona RG, et al. Effect of drug transporter genotypes on pravastatin disposition in European- and African-American participants. Pharmacogenet Genomics 2007;17(8):647–56.

[184] Pasanen MK, Neuvonen M, Neuvonen PJ, Niemi M. SLCO1B1 polymorphism markedly affects the pharmacokinetics of simvastatin acid. Pharmacogenet Genomics 2006;16(12):873–9.

[185] Donnelly LA, Doney AS, Tavendale R, Lang CC, Pearson ER, Colhoun HM, et al. Common nonsynonymous substitutions in SLCO1B1 predispose to statin intolerance in routinely treated individuals with type 2 diabetes: a go-DARTS study. Clin Pharmacol Ther 2011;89(2):210–6.

[186] Brunham LR, Lansberg PJ, Zhang L, Miao F, Carter C, Hovingh GK, et al. Differential effect of the rs4149056 variant in SLCO1B1 on myopathy associated with simvastatin and atorvastatin. Pharmacogenomics J 2012;12(3):233–7.

[187] Voora D, Shah SH, Spasojevic I, Ali S, Reed CR, Salisbury BA, et al. The SLCO1B1*5 genetic variant is associated with statin-induced side effects. J Am Coll Cardiol 2009;54(17):1609–16.

[188] Group SC, Link E, Parish S, Armitage J, Bowman L, Heath S, et al. SLCO1B1 variants and statin-induced myopathy–a genomewide study. N Engl J Med 2008;359(8):789–99.

[189] Hermann M, Bogsrud MP, Molden E, Asberg A, Mohebi BU, Ose L, et al. Exposure of atorvastatin is unchanged but lactone and acid metabolites are increased several-fold in patients with atorvastatin-induced myopathy. Clin Pharmacol Ther 2006;79(6):532–9.

[190] Danik JS, Chasman DI, MacFadyen JG, Nyberg F, Barratt BJ, Ridker PM. Lack of association between SLCO1B1 polymorphisms and clinical myalgia following rosuvastatin therapy. Am Heart J 2013;165(6):1008–14.

[191] Degorter MK, Tirona RG, Schwarz UI, Choi YH, Dresser GK, Suskin N, et al. Clinical and pharmacogenetic predictors of circulating atorvastatin and rosuvastatin concentrations in routine clinical care. Circ Cardiovasc Genet 2013;6(4):400–8.

[192] Mangravite LM, Engelhardt BE, Medina MW, Smith JD, Brown CD, Chasman DI, et al. A statin-dependent QTL for GATM expression is associated with statin-induced myopathy. Nature 2013;502(7471):377–80.

[193] Pirmohamed M, Friedmann PS, Molokhia M, Loke YK, Smith C, Phillips E, et al. Phenotype standardization for immune-mediated drug-induced skin injury. Clin Pharmacol Ther 2011;89(6):896–901.

[194] Alfirevic A, Pirmohamed M. Drug-induced hypersensitivity reactions and pharmacogenomics: past, present and future. Pharmacogenomics 2010;11(4):497–9.

[195] Chung W-H, Hung S-I, Chen Y-T. Human leukocyte antigens and drug hypersensitivity. Curr Opin Allergy Clin Immunol 2007;7(4):317–23. http://dx.doi.org/10.1097/ACI.0b013e3282370c5f.

[196] Pichler WJ, Naisbitt DJ, Park BK. Immune pathomechanism of drug hypersensitivity reactions. J Allergy Clin Immunol 2011;127(Suppl. 3):S74–81.

[197] Yun J, Adam J, Yerly D, Pichler WJ. Human leukocyte antigens (HLA) associated drug hypersensitivity: consequences of drug binding to HLA. Allergy 2012;67(11):1338–46.

[198] Hervey PS, Perry CM. Abacavir. Drugs 2000;60(2):447–79.

[199] Hetherington S, McGuirk S, Powell G, Cutrell A, Naderer O, Spreen B, et al. Hypersensitivity reactions during therapy with the nucleoside reverse transcriptase inhibitor abacavir. Clin Ther 2001;23(10):1603–14.

[200] Phillips EJ, Sullivan JR, Knowles SR, Shear NH. Utility of patch testing in patients with hypersensitivity syndromes associated with abacavir. AIDS 2002;16(16):2223–5.

[201] Peyrieere H, Nicolas J, Siffert M, Demoly P, Hillaire-Buys D, Reynes J. Hypersensitivity related to abacavir in two members of a family. Ann Pharmacother 2001;35(10):1291.

[202] Symonds W, Cutrell A, Edwards M, Steel H, Spreen B, Powell G, et al. Risk factor analysis of hypersensitivity reactions to abacavir. Clin Ther 2002;24(4):565–73.

[203] Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, et al. Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. Lancet 2002;359(9308):727–32.

[204] Hetherington S, Hughes AR, Mosteller M, Shortino D, Baker KL, Spreen W, et al. Genetic variations in HLA-B region and hypersensitivity reactions to abacavir. Lancet 2002;359(9312):1121–2.

[205] Hughes D, Vilar F, Ward C, Alfirevic A, Park BK, Pirmohamed M. Cost-effectiveness analysis of HLA B*5701 genotyping in preventing abacavir hypersensitivity. Pharmacogenetics 2004;14:335–42.

[206] Nolan D, Martin A, Almeida C, Mallal S. Prospective genetic screening decreases the incidence of abacavir hypersensitivity reactions in the Western Australian HIV cohort study. Clin Infect Dis 2006;1:99.

[207] Waters LJ, Mandalia S, Gazzard B, Nelson M. Prospective HLA-B*5701 screening and abacavir hypersensitivity: a single centre experience. AIDS 2007;21(18):2533–4.

[208] Zucman D, Majerholc C, Truchis PD, Stegman S, Caillat-Zucman S. Prospective screening for human leukocyte antigen-B*5701 avoids abacavir hypersensitivity reaction in the ethnically mixed French HIV population. J Acquir Immune Defic Syndr 2007;45(1):1–3.

[209] Colombo S, Rotger M, Martinez R, Telenti A, Rauch A, Fux C, et al. The HCP5 single-nucleotide polymorphism: a simple screening tool for prediction of hypersensitivity reaction to abacavir. J Infect Dis 2008;198(6):864–7.

[210] Dello Russo C, Lisi L, Lofaro A, Di Giambenedetto S, Federico B, Madeddu G, et al. Novel sensitive, specific and rapid pharmacogenomic test for the prediction of abacavir hypersensitivity reaction: HLA-B*57:01 detection by real-time PCR. Pharmacogenomics 2011;12(4):567–76.

[211] Mallal S, Phillips E, Nolan D, Carosi G, Molina JM, Workman C, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008;358(6):568–79.

[212] Schackman BR, Scott CA, Walensky RP, Losina E, Freedberg KA, Sax PE. The cost-effectiveness of HLA-B*5701 genetic screening to guide initial antiretroviral therapy for HIV. AIDS 2008;22(15):2025–33.

[213] Kauf TL, Farkouh RA, Earnshaw SR, Watson ME, Maroudas P, Chambers MG. Economic efficiency of genetic screening to inform the use of abacavir sulfate in the treatment of HIV. PharmacoEconomics 2010;28(11):1025–39.

[214] Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. Lancet 2007;369(9566):1000–15.

[215] Post RM, Ketter TA, Uhde T, Ballenger JC. Thirty years of clinical experience with carbamazepine in the treatment of bipolar illness: principles and practice. CNS Drugs 2007;21(1):47–71.

[216] Cruccu G, Gronseth G, Alksne J, Argoff C, Brainin M, Burchiel K, et al. AAN-EFNS guidelines on trigeminal neuralgia management. Eur J Neurol 2008;15(10):1013–28.

[217] Roujeau J-C, Kelly JP, Naldi L, Rzany B, Stern RS, Anderson T, et al. Medication use and the risk of Stevens-Johnson syndrome or toxic epidermal necrolysis. N Engl J Med 1995;333(24):1600–8.

[218] Borchers AT, Lee JL, Naguwa SM, Cheema GS, Gershwin ME. Stevens-Johnson syndrome and toxic epidermal necrolysis. Autoimmun Rev 2008;7(8):598–605.

[219] Mockenhaupt M, Viboud C, Dunant A, Naldi L, Halevy S, Bavinck JNB, et al. Stevens-Johnson syndrome and toxic epidermal necrolysis: assessment of medication risks with emphasis on recently marketed drugs. The EuroSCAR-study. J Invest Dermatol 2007;128(1):35–44.

[220] Schwartz RA, McDonough PH, Lee BW. Toxic epidermal necrolysis: part I. Introduction, history, classification, clinical features, systemic manifestations, etiology, and immunopathogenesis. J Am Acad Dermatol 2013;69(2):173.e1–13.

[221] Chung W-H, Hung S-I, Hong H-S, Hsih M-S, Yang L-C, Ho H-C, et al. Medical genetics: a marker for Stevens-Johnson syndrome. Nature 2004;428(6982):486.

[222] Man CBL, Kwan P, Baum L, Yu E, Lau KM, Cheng ASH, et al. Association between HLA-B*1502 allele and antiepileptic drug-induced cutaneous reactions in Han Chinese. Epilepsia 2007;48(5):1015–8.

[223] Wu XT, Hu FY, An DM, Yan B, Jiang X, Kwan P, et al. Association between carbamazepine-induced cutaneous adverse drug reactions and the HLA-B*1502 allele among patients in Central China. Epilepsy Behav 2010;19(3):405–8.

[224] Wang Q, Zhou J-Q, Zhou L-M, Chen Z-Y, Fang Z-Y, Chen S-D, et al. Association between HLA-B*1502 allele and carbamazepine-induced severe cutaneous adverse reactions in Han people of Southern China mainland. Seizure 2011;20(6):446–8.

[225] Tassaneeyakul W, Tiamkao S, Jantararoungtong T, Chen P, Lin S-Y, Chen W-H, et al. Association between HLA-B*1502 and carbamazepine-induced severe cutaneous adverse drug reactions in a Thai population. Epilepsia 2010;51(5):926–30.

[226] Locharernkul C, Loplumlert J, Limotai C, Korkij W, Desudchit T, Tongkobpetch S, et al. Carbamazepine and phenytoin induced Stevens-Johnson syndrome is associated with HLA-B*1502 allele in Thai population. Epilepsia 2008;49(12):2087–91.

[227] Then SM, Rani ZZM, Raymond AA, Ratnaningrum S, Jamal R. Frequency of the HLA-B*1502 allele contributing to carbamazepine-induced hypersensitivity reactions in a cohort of Malaysian epilepsy patients. Asian Pac J Allergy Immunol 2011;29(3):290–3.

[228] Mehta TY, Prajapati LM, Mittal B, Joshi CG, Sheth JJ, Patel DB, et al. Association of HLA-B*1502 allele and

carbamazepine-induced Stevens-Johnson syndrome among Indians. Indian J Dermatol Venereol Leprol 2009;75(6):579–82.

[229] Lim KS, Tan CT, Kwan P. Association of HLA-B*1502 allele and carbamazepine-induced severe adverse cutaneous drug reaction among Asians, a review. Neurol Asia 2008;13:15–21.

[230] Yip VL, Marson AG, Jorgensen AL, Pirmohamed M, Alfirevic A. HLA genotype and carbamazepine-induced cutaneous adverse drug reactions: a systematic review. Clin Pharmacol Ther 2012;92(6):757–65.

[231] Chen P, Lin J-J, Lu C-S, Ong C-T, Hsieh PF, Yang C-C, et al. Carbamazepine-induced toxic effects and HLA-B*1502 screening in Taiwan. N Engl J Med 2011;364(12):1126–33.

[232] Locharernkul C, Shotelersuk V, Hirankarn N. HLA-B* 1502 screening: time to clinical practice. Epilepsia 2010;51(5):936–8.

[233] Dong D, Sung C, Finkelstein EA. Cost-effectiveness of HLA-B 1502 genotyping in adult patients with newly diagnosed epilepsy in Singapore. Neurology 2012;79(12):1259.

[234] Ferrell PB, McLeod HL. Carbamazepine, HLA-B*1502 and risk of Stevens-Johnson syndrome and toxic epidermal necrolysis: US FDA recommendations. Pharmacogenomics 2008;9(10):1543–6.

[235] Kaniwa N, Saito Y, Aihara M, Matsunaga K, Tohkin M, Kurose K, et al. HLA-B locus in Japanese patients with anti-epileptics and allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis. Pharmacogenomics 2008;9(11):1617–22.

[236] Kim S-H, Lee KW, Song W-J, Kim S-H, Jee Y-K, Lee S-M, et al. Carbamazepine-induced severe cutaneous adverse reactions and HLA genotypes in Koreans. Epilepsy Res 2011;97(1–2):190–7.

[237] McCormack M, Alfirevic A, Bourgeois S, Farrell JJ, Kasperavičiūtė D, Carrington M, et al. HLA-A*3101 and carbamazepine-induced hypersensitivity reactions in Europeans. N Engl J Med 2011;364(12):1134–43.

[238] Ozeki T, Mushiroda T, Yowang A, Takahashi A, Kubo M, Shirakata Y, et al. Genome-wide association study identifies HLA-A*3101 allele as a genetic risk factor for carbamazepine-induced cutaneous adverse drug reactions in Japanese population. Hum Mol Genet 2011;20(5):1034–41.

[239] Hung S-I, Chung W-H, Jee S-H, Chen W-C, Chang Y-T, Lee W-R, et al. Genetic susceptibility to carbamazepine-induced cutaneous adverse drug reactions. Pharmacogenet Genomics 2006;16(4):297–306.

[240] Pacher P, Nivorozhkin A, Szabó C. Therapeutic effects of xanthine oxidase inhibitors: renaissance half a century after the discovery of allopurinol. Pharmacol Rev 2006;58(1):87–114.

[241] Hung S-I, Chung W-H, Liou L-B, Chu C-C, Lin M, Huang H-P, et al. HLA-B*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. Proc Nat Acad Sci USA 2005;102(11):4134–9.

[242] Cao Z-H, Wei Z-Y, Zhu Q-Y, Zhang J-Y, Yang L, Qin S-Y, et al. HLA-B*58:01 allele is associated with augmented risk for both mild and severe cutaneous adverse reactions induced by allopurinol in Han Chinese. Pharmacogenomics 2012;13(10):1193–201.

[243] Tassaneeyakul W, Jantararoungtong T, Vannaprasaht S, Tiamkao S, Choonhakarn C, Tassaneeyakul W, et al. Strong association between HLA-B*5801 and allopurinol-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in a Thai population. Pharmacogenet Genomics 2009;19(9):704–9.

[244] Tohkin M, Kaniwa N, Saito Y, Sugiyama E, Kurose K, Nishikawa J, et al. A whole-genome association study of major determinants for allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis in Japanese patients. Pharmacogenomics J 2013;13(1):60–9.

[245] Kang HR, Kim YS, Jung JW, Kim SH, Park HW, Chang YS, et al. Positive and negative associations of HLA class I alleles with allopurinol-induced SCARs in Koreans. Pharmacogenet Genomics 2011;21(5):303–7.

[246] Lonjou C, Borot N, Ledger N, Hovnanian A, de Toma C, Roujeau JC, et al. A European study of HLA-B in Stevens-Johnson syndrome and toxic epidermal necrolysis related to five high-risk drugs. Pharmacogenet Genomics 2008;18(2):99–107.

[247] Bigby M. Rates of cutaneous reactions to drugs. Arch Dermatol 2001;137(6):765–70.

[248] Yawalkar N. Drug-induced exanthems. Toxicology 2005;209(2):131–4.

[249] Forget E, Menzies D. Adverse reactions to first-line antituberculosis drugs. Expert Opin Drug Saf 2006;5(2):231–49.

[250] American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America. Treatment of tuberculosis. Am J Respir Crit Care Med 2003;167(4):603–62.

[251] Kim S-H, Kim S-H, Yoon H, Shin D, Park S, Kim Y-S, et al. NAT2, CYP2C9, CYP2C19, and CYP2E1 genetic polymorphisms in anti-TB drug-induced maculopapular eruption. Eur J Clin Pharmacol 2011;67(2):121–7.

[252] Phillips KA, Veenstra DL, Oren E, Lee JK, Sadee W. Potential role of pharmacogenomics in reducing adverse drug reactions: a systematic review. JAMA 2001;286(18):2270–9.

[253] Goldstein JA, Ishizaki T, Chiba K, de Morais SMF, Bell D, Krahn PM, et al. Frequencies of the defective CYP2C19 alleles responsible for the mephenytoin poor metabolizer phenotype in various oriental, Caucasian, Saudi Arabian and American black populations. Pharmacogenetics 1997;7(1):59.

[254] Roujeau JC, Stern RS. Severe adverse cutaneous reactions to drugs. N Engl J Med 1994;331(19):1272–85.

[255] McAuley JW, Biederman TS, Smith JC, Moore JL. Newer therapies in the drug treatment of epilepsy. Ann Pharmacother 2002;36(1):119–29.

[256] Garcia-Doval I, LeCleach L, Bocquet H, Otero X, Roujeau J. Toxic epidermal necrolysis and stevens-johnson syndrome: does early withdrawal of causative drugs decrease the risk of death? Arch Dermatol 2000;136(3):323–7.

[257] Hu F-Y, Wu X-T, An D-M, Yan B, Stefan H, Zhou D. Pilot association study of oxcarbazepine-induced mild cutaneous adverse reactions with HLA-B*1502 allele in Chinese Han population. Seizure 2011;20(2):160–2.

[258] Li L-J, Hu F-Y, Wu X-T, An D-M, Yan B, Zhou D. Predictive markers for carbamazepine and lamotrigine-induced maculopapular exanthema in Han Chinese. Epilepsy Res 2013;106(1-2):296–300.

[259] He N, Min F-L, Shi Y-W, Guo J, Liu X-R, Li B-M, et al. Cutaneous reactions induced by oxcarbazepine in Southern Han Chinese: incidence, features, risk factors and relation to HLA-B alleles. Seizure 2012;21(8):614–8.

[260] Arif H, Buchsbaum R, Weintraub D, Koyfman S, Salas-Humara C, Bazil CW, et al. Comparison and predictors of rash associated with 15 antiepileptic drugs. Neurology 2007;68(20):1701–9.

[261] An D-M, Wu X-T, Hu F-Y, Yan B, Stefan H, Zhou D. Association study of lamotrigine-induced cutaneous adverse reactions and HLA-B*1502 in a Han Chinese population. Epilepsy Res 2010;92(2–3):226–30.

[262] Shi Y-W, Min F-L, Liu X-R, Zan L-X, Gao M-M, Yu M-J, et al. HLA-B alleles and lamotrigine-induced cutaneous adverse drug reactions in the Han Chinese population. Basic Clin Pharmacol Toxicol 2011;109(1):42–6.

[263] Schachter SC. Oxcarbazepine: current status and clinical applications. Expert Opin Investig Drugs 1999;8(7):1103.

[264] Alvestad S, Lydersen S, Brodtkorb E. Rash from antiepileptic drugs: influence by gender, age, and learning disability. Epilepsia 2007;48(7):1360–5.

[265] Wang X-Q, Lang S-Y, Shi X-B, Tian H-J, Wang R-F, Yang F. Antiepileptic drug-induced skin reactions: a retrospective study and analysis in 3793 Chinese patients with epilepsy. Clin Neurol Neurosurg 2012;114(7):862–5.

[266] Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, et al. Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89(6):806–15.

[267] Temple R. Hepatotoxicity through the years, www.fda.gov/downloads/Drugs/ScienceResearch/ResearchAreas/ucm122149.pdf; 2006 [accessed 23.01.13].

[268] Ostapowicz G, Fontana RJ, Schiødt FV, Larson A, Davern TJ, Han SHB, et al. Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. Ann Intern Med 2002;137(12):947–54.

[269] Sgro C, Clinard F, Ouazir K, Chanay H, Allard C, Guilleminet C, et al. Incidence of drug-induced hepatic injuries: a French population-based study. Hepatology 2002;36(2):451–5.

[270] Björnsson E, Jerlstad P, Olsson R, Bergqvist A. Fulminant drug-induced hepatic failure leading to death or liver transplantation in Sweden. Scand J Gastroenterol 2005;40(9):1095–101.

[271] Russmann S, Kullak-Ublick GA, Grattagliano I. Current concepts of mechanisms in drug-induced hepatotoxicity. Curr Med Chem 2009;16(23):3041–53.

[272] Russmann S, Jetter A, Kullak-Ublick GA. Pharmacogenetics of drug-induced liver injury. Hepatology 2010;52(2):748–61.

[273] Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenomics J 2008;8(3):186–95.

[274] Singer JB, Lewitzky S, Leroy E, Yang F, Zhao X, Klickstein L, et al. A genome-wide study identifies HLA alleles associated with lumiracoxib-related liver injury. Nat Genet 2010;42(8):711–4.

[275] Andrade RJ, García-Ruiz E, García-Muñoz B, González-Grande R, Pizarro A, Lucena MI, et al. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. Gastroenterology 2005;129(2):512–21.

[276] Chalasani N, Fontana RJ, Bonkovsky HL, Watkins PB, Davern T, Serrano J, et al. Causes, clinical features, and outcomes from a prospective study of drug-induced liver injury in the United States. Gastroenterology 2008;135(6):1924–34, 1934.e1–4.

[277] García Rodríguez LA, Stricker BH, Zimmerman HJ. Risk of acute liver injury associated with the combination of amoxicillin and clavulanic acid. Arch Intern Med 1996;156(12):1327–32.

[278] Fontana RJ, Shakil AO, Greenson JK, Boyd I, Lee WM. Acute liver failure due to amoxicillin and amoxicillin/clavulanate. Dig Dis Sci 2005;50(10):1785–90.

[279] Hautekeete ML, Horsmans Y, van Waeyenberge C, Demanet C, Henrion J, Verbist L, et al. HLA association of amoxicillin-clavulanate-induced hepatitis. Gastroenterology 1999;117(5):1181–6.

[280] O'Donohue J, Mills PR, Oien KA, MacSween RNM, Donaldson P, Underhill J, et al. Co-amoxiclav jaundice: clinical and histological features and HLA class II association. Gut 2000;47(5):717–20.

[281] Donaldson PT, Daly AK, Henderson J, Graham J, Pirmohamed M, Bernal W, et al. Human leukocyte antigen class II genotype in susceptibility and resistance to co-amoxiclav-induced liver injury. J Hepatol 2010;53(6):1049–53.

[282] Andrade RJ, Lucena MI, Alonso A, García-Cortes M, García-Ruiz E, Benitez R, et al. HLA class II genotype influences the type of liver injury in drug-induced idiosyncratic liver disease. Hepatology 2004;39(6):1603–12.

[283] Lucena MI, Andrade RJ, Stephens C, Molokhia M, Shen Y, Floratos A, et al. Susceptibility to amoxicillin-clavulanate-induced liver injury is influenced by multiple HLA class I and II alleles. Gastroenterology 2011;141(1):338–47.

[284] Russmann S, Kaye JA, Jick SS, Jick H. Risk of cholestatic liver disease associated with flucloxacillin and flucloxacillin prescribing habits in the UK: cohort study using data from the UK general practice research database. Br J Clin Pharmacol 2005;60(1):76–82.

[285] Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe'er I, Floratos A, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. Nat Genet 2009;41(7):816–9.

[286] Alfirevic A, Pirmohamed M. Predictive genetic testing for drug-induced liver injury: considerations of clinical utility. Clin Pharmacol Ther 2012;92(3):376–80.

[287] Testa L, Bhindi R, van Gaal WJ, Agostoni P, Abbate A, Biondi Zoccai GGL. The direct thrombin inhibitor ximelagatran/melagatran: a systematic review on clinical applications and an evidence based assessment of risk benefit profile. Expert Opin Drug Saf 2007;6(4):397–406.

[288] Gibert M, Sanchez-Mazas A. Geographic patterns of functional categories of HLA-DRB1 alleles: a new approach to analyse associations between HLA-DRB1 and disease. Eur J Immunogenet 2003;30(5):361–74.

[289] Harenberg J, Jörg I, Weiss C. Observations of alanine aminotransferase and aspartate aminotransferase in THRIVE studies treated orally with ximelagatran. Int J Toxicol 2006;25(3):165–9.

[290] Gøtzsche P, Jørgensen K. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013;6:CD001877.

[291] Martin AM, Nolan D, James I, Cameron P, Keller J, Moore C, et al. Predisposition to nevirapine hypersensitivity associated with HLA-DRB1 0101 and abrogated by low CD4 T-cell counts. AIDS 2005;19(1):97.

[292] Yuan J, Guo S, Huang Z, Hall D, Cammett AM, Storfer S, et al. Toxicogenomics of nevirapine-associated cutaneous and hepatic adverse events among populations of African, Asian, and European descent. AIDS 2011;25(10):1271–80.

[293] Spraggs CF, Budde LR, Briley LP, Bing N, Cox CJ, King KS, et al. HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. J Clin Oncol 2011;29(6):667–73.

[294] Hirata K, Yamamoto M, Okutani Y, Takagi H, Matsumoto T, Nishiya T, et al. Ticlopidine-induced hepatotoxicity is associated with specific human leukocyte antigen genomic subtypes in Japanese patients: a preliminary case-control study. Pharmacogenomics J 2008;8(1):29–33.

[295] Ariyoshi N, Iga Y, Hirata K, Sato Y, Miura G, Ishii I, et al. Enhanced susceptibility of HLA-mediated ticlopidine-induced idiosyncratic

hepatotoxicity by CYP2B6 polymorphism in Japanese. Drug Metab Pharmacokinet 2010;25(3):298.

[296] Shubber Z, Calmy A, Andrieux-Meyer I, Vitoria M, Renaud-Théry F, Shaffer N, et al. Adverse events associated with nevirapine and efavirenz-based first-line antiretroviral therapy: a systematic review and meta-analysis. AIDS 2013;27(9):1403–12.

[297] Goss PE, Smith IE, O'Shaughnessy J, Ejlertsen B, Kaufmann M, Boyle F, et al. Adjuvant lapatinib for women with early-stage HER2-positive breast cancer: a randomised, controlled, phase 3 trial. Lancet Oncol 2013;14(1):88–96.

[298] Mizushima M, Iwata N, Fujimoto T-T, Ishikawa K, Fujimura K. Patient characteristics in ticlopidine hydrochloride-induced liver injury: case–control study. Hepatol Res 2005;33(3):234–40.

[299] Bonierbale E, Valadon P, Pons C, Desfosses B, Dansette PM, Mansuy D. Opposite behaviors of reactive metabolites of tienilic acid and its isomer toward liver proteins: use of specific anti-tienilic acid–protein adduct antibodies and the possible relationship with different hepatotoxic effects of the two compounds. Chem Res Toxicol 1999;12(3):286.

[300] Pirmohamed M. Acceptance of biomarker-based tests for application in clinical practice: criteria and obstacles. Clin Pharmacol Ther 2010;88(6):862–6.

[301] Aithal GP, Rawlins MD, Day CP. Accuracy of hepatic adverse drug reaction reporting in one English health region. BMJ 1999;319(7224):1541.

[302] Rockey DC, Seeff LB, Rochon J, Freston J, Chalasani N, Bonacini M, et al. Causality assessment in drug-induced liver injury using a structured expert opinion process: comparison to the Roussel-Uclaf causality assessment method. Hepatology 2010;51(6):2117–26.

[303] Brugts JJ, Isaacs A, Boersma E, van Duijn CM, Uitterlinden AG, Remme W, et al. Genetic determinants of treatment benefit of the angiotensin-converting enzyme-inhibitor perindopril in patients with stable coronary artery disease. Eur Heart J 2010;31(15):1854.

[304] Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 2008;40(6):695–701.

[305] Szyf M. The dynamic epigenome and its implications in toxicology. Toxicol Sci 2007;100(1):7–23.

[306] Kacevska M, Ivanov M, Ingelman-Sundberg M. Perspectives on epigenetics and its relevance to adverse drug reactions. Clin Pharmacol Ther 2011;89(6):902–7.

[307] Veenstra D, Higashi M, Phillips K. Assessing the cost-effectiveness of pharmacogenomics. AAPS PharmSci 2000;2(3):80–90.

[308] Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther 2008;84(3):362–9.

[309] Ross CJD, Visscher H, Sistonen J, Brunham LR, Pussegoda K, Loo TT, et al. The Canadian pharmacogenomics network for drug safety: a model for safety pharmacology. Thyroid 2010;20(7):681.

This page is intentionally left blank

Chapter 21

HLA and the Pharmacogenomics of Drug Hypersensitivity

Katie D. White,[a] **Silvana Gaudieri,**[b,c] **and Elizabeth J. Phillips**[a,c]

[a]*Department of Medicine, Vanderbilt University Medical Center, USA,* [b]*School of Anatomy, Physiology and Human Biology, University of Western Australia, Australia,* [c]*Institute for Immunology and Infectious Diseases, Murdoch University, Australia*

Chapter Outline

21.1 Overview 437
21.2 Introduction 437
  21.2.1 MHC Class I 438
  21.2.2 MHC Class II 439
  21.2.3 HLA Diversity, Linkage Disequilibrium, and Specificity of Peptide Binding 439
21.3 Classification of Drug Hypersensitivity Syndromes 440
  21.3.1 DIHS/DRESS/HSS 440
  21.3.2 SJS/TEN 441
  21.3.3 Single-Organ Drug Hypersensitivity Syndromes 441
21.4 Immunopathogenesis of Drug Hypersensitivity Syndromes 442
  21.4.1 The Hapten/Prohapten Model 442
  21.4.2 The p-i Model 443
  21.4.3 The Altered Peptide Repertoire Model 444
21.5 HLA and Hypersensitivity Reactions 444
  21.5.1 Abacavir 444
  21.5.2 Carbamazepine 450
  21.5.3 Allopurinol 452
  21.5.4 Nevirapine 453
  21.5.5 Amoxicillin-Clavulanate 453
21.6 Stratified Approaches to Prevent Drug Hypersensitivity and Their Translation into Clinical Medicine 455
  21.6.1 Criteria for Stratified Approaches 455
  21.6.2 Current Recommendations for HLA Screening 457
21.7 The Future Direction of Stratified Medicine in Drug-Induced Hypersensitivity 458
Acronyms 459
References 459

21.1 OVERVIEW

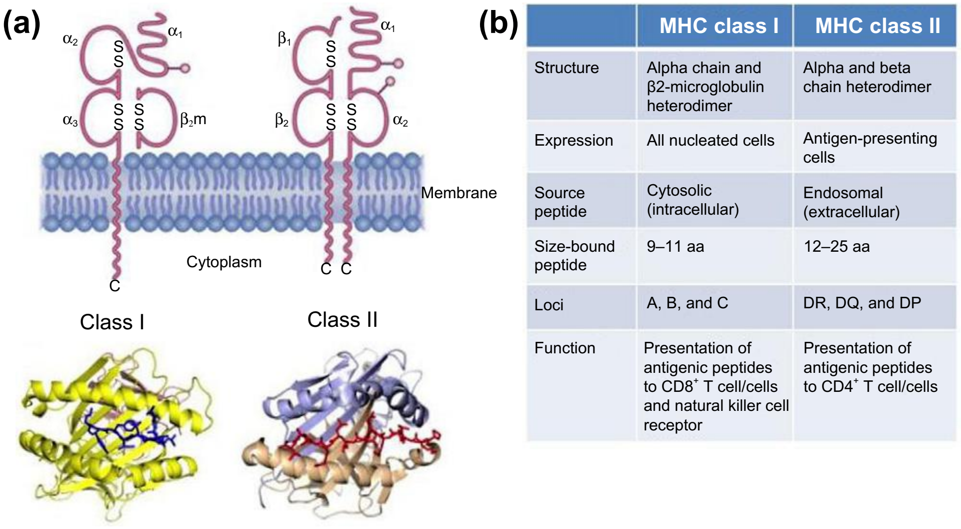
Adverse drug reactions (ADRs) are a source of tremendous economic burden to the healthcare community and are common causes of patient morbidity and mortality, estimated to be the fourth to sixth most common causes of death in some series [1–4]. In addition, ADRs are responsible for approximately 3–6% of inpatient admissions and comprise 5–10% of inpatient costs, with expenses approximating $4 billion per year in the United States alone. Although fatal ADRs occur in less than 1% of hospitalized patients, it is estimated that at least 50% of these are preventable [5–8].

ADRs are classified into two categories. Type A ADRs comprise 80% of reported drug reactions and are relatively predictable based on the pharmacologic activity of the drug. Type B ADRs account for <20% of reported drug reactions and are less dependent on dose and not readily anticipated based on pharmacologic principles. It is now appreciated that most Type B reactions are primarily immunologically mediated and that many seemingly idiosyncratic drug reactions may be predicted based on patient immunogenetics [9,10]. This chapter explores our current understanding of the drug reactions that are secondary to T-cell immune responses and the association with particular human leukocyte antigen (HLA) alleles.

21.2 INTRODUCTION

Recent work has established that the HLA family of proteins, encoded within the major histocompatibility complex (MHC), is central to the immunopathogenesis of many

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00021-9
© 2014 Elsevier Inc. All rights reserved.



	MHC class I	MHC class II
Structure	Alpha chain and β2-microglobulin heterodimer	Alpha and beta chain heterodimer
Expression	All nucleated cells	Antigen-presenting cells
Source peptide	Cytosolic (intracellular)	Endosomal (extracellular)
Size-bound peptide	9–11 aa	12–25 aa
Loci	A, B, and C	DR, DQ, and DP
Function	Presentation of antigenic peptides to CD8$^+$ T cell/cells and natural killer cell receptor	Presentation of antigenic peptides to CD4$^+$ T cell/cells

FIGURE 21.1 **HLA structure and peptide binding pockets.** (a) Schematic of the MHC class I and II molecules and their respective binding pockets. The MHC class I molecule is a heterodimer containing the transmembrane α domains and β_2-microglobulin. The α1 and α2 domains form the peptide binding pocket that can accommodate peptides 9–11 amino acids in length. The MHC class II is a heterodimer composed of two polypeptide chains, α and β. The peptide binding pocket is formed by α1 and β1. The open-ended pocket can accommodate peptides of 12–25 amino acids in length. (b) Differences between MHC class I and class I molecules. *Source: Adapted from: www.uptomed.ir/Digimed.ir/cecil/Cecil/HTML/188.htm and http://nfs.unipv.it/nfs/minf/dispense/immunology/lectures/files/structure_mhc.html.*

Type B ADRs. Broadly grouped into two structurally and functionally distinct classes, HLA class I and HLA class II (also known as MHC class I and MHC class II), these proteins bind and present peptide antigens to T cells and thus are central to immune surveillance (Figure 21.1). MHC class I molecules are expressed by most nucleated cells and generally serve to display peptides that are derived from intracellular proteins (e.g., misfolded self-proteins and viral proteins) to cytotoxic CD8+ T cells. The expression of MHC class II proteins is restricted to specialized antigen-presenting cells (APCs) (e.g., dendritic cells and monocytes) that interact with CD4+ helper or regulatory T cells [9,11]. All genes encoding the classical MHC proteins are located on chromosome 6—*HLA-A*, *HLA-B*, and *HLA-C* for class I MHC and *HLA-DR*, *HLA-DQ*, and *HLA-DP* for class II MHC proteins. The HLA genes are the most polymorphic genetic elements in the human genome and as a result there exists an extraordinary amount of diversity among MHC proteins [12,13].

21.2.1 MHC Class I

The MHC class I molecule (MHCI) is a heterodimeric protein composed of a transmembrane α heavy chain and a smaller, extracellular β_2-microglobulin polypeptide (Figure 21.1). The α-chain is folded into three distinct domains, α1, α2 and α3. The conserved, membrane-proximal α3 and β_2-microglobulin domains each assume an immunoglobulin-like conformation with the α3 domain serving an anchor to the cell surface. The N-terminal α1α2 domains possess a high degree of polymorphism, and each contributes four antiparallel β strands and an α-helical motif to form the floor and walls of the MHCI peptide binding groove [14]. Within the MHCI binding cleft, chemically distinct binding pockets (denoted pockets A–F) interact with the amino acid side chains of bound peptide thereby conferring specificity to the selection of peptide antigen. Peptide amino acid side chains that are not anchored to the MHCI binding cleft remain exposed and accessible to interaction with the T-cell receptor (TCR) [15]. Importantly, the fidelity of the MHCI interaction with CD8+ T cells (and not APCs displaying CD4) is maintained by CD8 binding to invariant regions of the MHCI molecule [11].

Association of peptide antigen with MHCI occurs in the endoplasmic reticulum (ER) following proteasome-mediated degradation of cellular (either defective self- or pathogen-associated) proteins into short peptide fragments. Resultant peptides are generally 9–11 amino acids in length and are shuttled into the ER via the integral membrane protein TAP (transporter associated with antigen processing). TAP and the nascent MHCI protein then associate with specific chaperone proteins (e.g., ERp57 and tapasin) as well as peptide fragments to form the peptide-loading complex (PLC). The peptide-loaded MHCI complex then traffics to and is expressed on the surface of the cell membrane, primed for contact with T cells. MHCI-peptide binding occurs at nanomolar affinities, thus ensuring that a long-lived complex is available at the cell surface for T-cell surveillance [16–18].

The process of T-cell selection in the thymus generates a pool of CD8+ T cells that interact weakly with self antigens presented on MHCI, thereby failing to elicit a fulminant immune response to self-proteins. If, however, a foreign peptide such as that derived from a viral protein is presented in the context of MHCI, the TCR will bind with high affinity and, with appropriate concomitant costimulatory signals, will elicit a robust cytotoxic response that results in cell-mediated killing of the target cell. In this way, the MHCI protein serves to continuously sample and present intracellular antigens, allowing the CD8+ T-cell repertoire to survey the intracellular peptide pool and destroy cells that are producing “foreign” proteins [9,19].

21.2.2 MHC Class II

The MHC class II molecule (MHCII) is composed of two polypeptide chains, α and β, each of which traverses the cell membrane and associates to form a heterodimer containing two conserved immunoglobulin-like domains proximal to the cell membrane (α2β2) and two polymorphic domains that associate to form the peptide binding site (α1β1) (Figure 21.1). In contrast to MHCI, the peptide binding cleft of the MHCII protein is open-ended and can therefore accommodate larger peptide fragments, generally 12–25 amino acids in length. This feature results in the selection of a broader range of peptides bound and presented in the context MHCII and makes it difficult to predict which peptides will bind to an MHCII molecule based on sequence alone. Structural analysis of the HLA-DR1 protein demonstrates that the peptide binding groove contains a hydrophobic pocket that creates a binding site for the side chain of the “anchoring” peptide residue, with additional hydrogen bonds occurring between the extended peptide main chain and conserved residues of the MHCII binding groove α-helices [20]. These conserved residues in the peptide binding groove are interspersed with variable amino acids that interact with peptide side-chain moieties and thus provide specificity to the peptide pool presented [20–22].

Antigens presented in the context of MHCII are derived from the extracellular environment and are processed into peptide fragments by endosomal machinery following endocytosis. As mentioned previously, the expression of MHCII is restricted to APCs (dendritic cells, macrophages, and B lymphocytes), and peptide displayed in the context of MHCII is restricted to CD4+ T-cell surveillance via CD4–MHCII binding. The immune response elicited following stimulation of the CD4+ T cell by the MHCII/peptide complex is variable and distinct from that elicited by the MHCI/CD8+ T-cell interaction.

21.2.3 HLA Diversity, Linkage Disequilibrium, and Specificity of Peptide Binding

The HLA genes are extraordinarily diverse, with greater than 1000 allelic variants described for the Class I HLA-A, HLA-B, and HLA-C and Class II HLA-DRB1 loci [13] (Figure 21.2). The majority of HLA polymorphism occurs among the peptide binding residues such that each HLA allotype presents a unique repertoire of peptide ligands for T-cell surveillance. Presumably, this high degree of polymorphism is the result of natural selection for protective immunity against infectious pathogens [23,24]. Heterozygosity at the HLA loci theoretically results in a broader repertoire of presented peptides, thereby increasing the probability that a pathogen will be recognized and that an immune response will be elicited [12,13,25–27]. However, the potential overall diversity across the different HLA loci that could be present within a population is reduced by the linkage disequilibrium (LD) observed within and between the HLA class I and HLA class II loci [28]. This results in the formation of MHC haplotypes (also referred to as ancestral haplotypes) that can span megabases of sequence. The LD is greatest between more closely linked HLA molecules (e.g., HLA-B and HLA-C or HLA-DRB1 and HLA-DQB1) but can also include combinations of HLA class I and II genes. The selective advantages conferred by the maintenance of these MHC haplotypes through human history remain unclear. In the context of disease association studies, attempting to disentangle associations with closely linked HLA genes can be difficult and often requires recombinant mapping and genome-wide association studies (GWAS) using single-nucleotide polymorphisms (SNPs) to tag regions of high LD that include one or more HLA loci.

Clinical and laboratory studies have demonstrated that even minor changes in the amino acid sequence of the MHC-peptide binding groove can significantly impact the presentation of antigen and the resultant immune response. For example, Epstein-Barr viral peptides displayed in the context of the HLA-B44 allotypes HLA-B*44:02, HLA-B*44:03, and HLA-B*44:05, which differ from each other at only two amino acids in the peptide binding groove, assume subtle conformational changes that result in marked differences in the affinity of the TCR-MHCI/peptide interaction [29]. Further, the differential peptide selection between the HLA-B*44:02 and HLA-B*44:03 allotypes has been shown to pose a significant barrier to hematopoietic stem cell transplantation with a high rate of acute allograft rejection identified in B*44:02/B*44:03 mismatched donor-recipient pairs [15]. It has also been shown that a single amino acid difference in the peptide binding groove between individuals expressing the HLA-B*35:01 and HLA-B*35:08 alleles results in the selection of distinct sets of immunodominant viral peptides during natural infection with human cytomegalovirus [30]. Similarly, in HIV-infected patients, single amino acid differences in the HLA-B*35:01 allele can influence the rate of progression to AIDS [31]. Such heterologous immune responses also play a significant role in solid-organ transplant rejection and likely in other inflammatory diseases, including drug hypersensitivity [32,33].

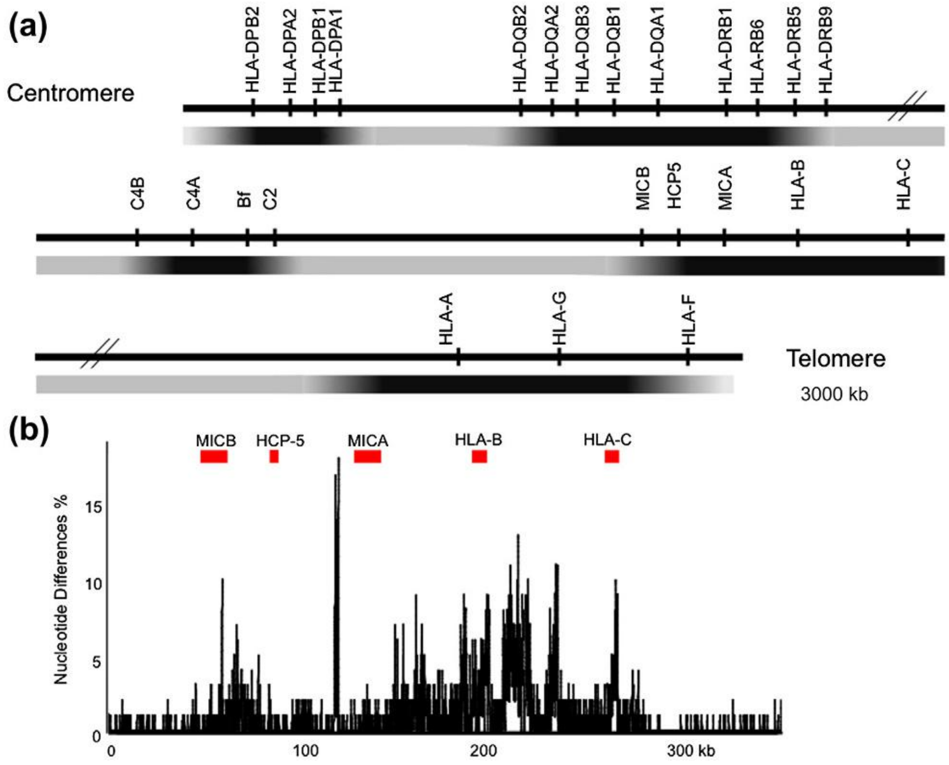


FIGURE 21.2 **Structure of the Major Histocompatibility Complex (MHC).** (a) Map of the MHC including classical class I and II loci and central MHC genes (e.g., MICA, MICB) complement and HCP5. The intensity of the box underneath the different regions reflects the LD observed between different loci (the darker the intensity, the greater the LD). Note the LD from the centromeric to the telomeric end of the MHC covering approximately 3000 kilobases (kb) of sequence [28]. (b) Nucleotide diversity plot between two MHC haplotypes in a 350 kb region containing MICB, MICA, HCP5, HLA-B, and HLA-C [35]. High levels of diversity are associated with low levels of recombination. Accordingly, the region between HLA-C and MICB exhibits high LD (the level of diversity observed here is more than 10-fold elsewhere in the genome [35]). However, note that between HCP5 and HLA-B are areas of diminished diversity that may correspond to the potential recombination spots. The amount of recombination in an area is inversely proportional to the ability of a marker to tag a nearby gene. In this region, recombination is not common and SNPs can act as haplotypic or haplospecific markers for a nearby gene. For example, variations in HCP5 and HLA-B*57:01 are in high LD, but there are individuals in which this is not the case [36], likely representing historical recombination events.

These examples demonstrate that minor differences in the amino acid composition of the peptide binding groove—termed micropolymorphisms—have the capacity to dramatically alter peptide selection and the immune response to viral or foreign-tissue antigens. These concepts translate into the variable susceptibility seen in individuals of differing HLA allotypes to ADRs as described next.

21.3 CLASSIFICATION OF DRUG HYPERSENSITIVITY SYNDROMES

The drug hypersensitivity syndromes (DHS) encompass phenotypically distinct clinical diagnoses. Traditionally, those that are T-cell mediated have been classified as Gell-Coombs type IV reactions, and this classification scheme has recently been modified to include four subtypes to represent the effector cell involved (subtypes a–d; Table 21.1) [34]. The syndromes include drug-induced hypersensitivity syndrome (DIHS)/drug reaction with eosinophilia (DRESS)/hypersensitivity syndrome (HSS), Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN), acute generalized exanthematous pustulosis (AGEP), drug-induced lupus, fixed drug eruption, and single-organ involvement pathologies such as drug-induced liver disease (DILI). With the exception of DILI, where toxic mechanisms may be important, all of these syndromes are thought to be primarily immunologically mediated. Since many of these DHS are associated with significant cutaneous disease, the term *severe cutaneous adverse reaction* (SCAR) has also been used to collectively refer to DIHS/DRESS, SJS/TEN, and AGEP.

21.3.1 DIHS/DRESS/HSS

DIHS presents as a widespread rash of varying severity, but without skin separation or blistering, accompanied by a minimum criteria of fever, internal organ involvement (usually hepatitis), and hematologic abnormalities (often atypical lymphocytes

TABLE 21.1 Revised Gell and Coombs Classification of Hypersensitivity Reactions

Type of Reaction	Antigen	Immune Reactant	Effector	Clinical Symptoms
Type I (immediate)	Soluble antigen	IgE	Mast cell, basophil	Pruritus, angioedema, urticaria, bronchospasm
Type II (cytotoxic)	Cell- or matrix-associated antigen	IgG	FcR positive cells (phagocytes, NK cells)	Thrombocytopenia, hemolytic anemia
Type III (immune complex)	Soluble antigen	IgG	FcR positive cells, complement	Serum sickness
Type IV (delayed)	Antigen presented by HLA proteins or direct T-cell stimulation	T-cell–mediated		
		IVa	Macrophage	Contact dermatitis
		IVb	Eosinophils	DRESS/DIHS/HSS
		IVc	T cells	SJS/TEN, DILI
		IVd	Neutrophils	AGEP

Source: Adapted from Rive et al. [34].

Ig—immunoglobulin; FcR—Fc receptor; NK cells—natural killer cells; DRESS—drug reaction with eosinophilia and systemic symptoms; DIHS—drug-induced hypersensitivity syndrome; HSS—hypersensitivity syndrome; SJS—Stevens-Johnson syndrome; TEN—toxic epidermal necrolysis; DILI—drug-induced liver injury; AGEP—acute generalized exanthematous pustulosis.

and/or eosinophilia). Diffuse lymphadenopathy, pneumonitis, encephalitis, cardiac failure (myocarditis), and nephritis are variable features of this syndrome, which may mimic a viral illness. Onset of symptoms typically occurs 2–8 weeks following initiation of the inciting drug and can persist for weeks [37]. Prolonged or recurrent symptoms, sometimes weeks following cessation of the offending drug, have been linked to reactivation of the human herpes viruses (HHV-6, CMV) [38,39]. Numerous drugs have been shown to be associated with the development of DIHS, including the antiretrovirals (abacavir, nevirapine, fosamprenavir), allopurinol, antiepileptic medications (carbamazepine, phenytoin, phenobarbital, lamotrigine), beta-lactam antibiotics, NSAIDs, and sulfa antimicrobials. Abacavir does not cause typical DIHS/DRESS/HSS in that it is not characteristically associated with hematological abnormalities or internal organ involvement; rash, which is usually mild to moderate, is a late and inconsistent feature.

21.3.2 SJS/TEN

SJS and TEN are the most severe cutaneous hypersensitivity syndromes and are thought to be a spectrum of disease defined by the percentage of total body surface area (TBSA) involvement. They are characterized by a painful blistering skin rash that can result in significant epidermal necrosis and detachment, sepsis, and death. These syndromes are often associated with multi-organ involvement (anemia, lymphopenia, complications from epidermal fluid losses, and superimposed bacterial infection/sepsis). Fever is common, and patients may complain of sore throat and/or visual impairment as a result of oropharyngeal mucous membrane and conjunctival erosions. Generally, SJS is used to define cases where blistering and epidermal detachment occur on <10% of TBSA; TEN is diagnosed in instances where >30% of TBSA is affected. SJS/TEN describes the overlap in this spectrum in which 10–30% of TBSA is affected. However, patients with severe epidermal detachment and skin necrosis are classified as TEN regardless of the percentage of TBSA involved.

TEN and SJS are often associated with significant mortality as well as short- and long-term morbidity, including respiratory, gastrointestinal, and genitourinary complications as well as permanent corneal scarring and vision loss [1,40]. The mortality associated with TEN may reach 30–50% [40]. Drugs implicated in the pathogenesis of SJS/TEN include the sulfa antimicrobials (antimicrobial sulfonamides), the antiretroviral drug nevirapine, allopurinol, lamotrigine, the aromatic anticonvulsants, and NSAIDs [10]. The mainstay of treatment is the prompt discontinuation of the inciting medication and supportive therapies, often in a specialized burn or intensive care unit. Rechallenge with an offending drug induces rapid recurrence of symptoms and, given the morbidity and mortality associated with SJS/TEN, permanent avoidance of all suspected offending drugs and structurally related drugs is recommended.

21.3.3 Single-Organ Drug Hypersensitivity Syndromes

Isolated hepatic dysfunction has been associated with the antimicrobial drugs flucloxacillin, amoxicillin-clavulanate

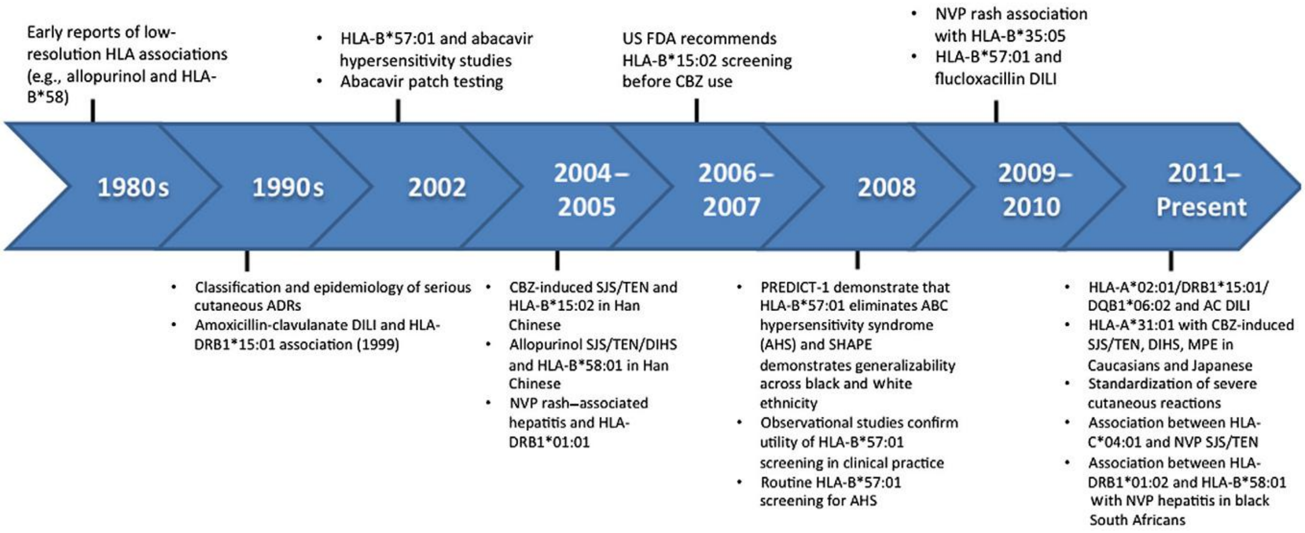

Early reports of low-resolution HLA associations (e.g., allopurinol and HLA-B*58)
• HLA-B*57:01 and abacavir hypersensitivity studies
• Abacavir patch testing
US FDA recommends HLA-B*15:02 screening before CBZ use
• NVP rash association with HLA-B*35:05
• HLA-B*57:01 and flucloxacillin DILI
1980s
1990s
2002
2004–2005
2006–2007
2008
2009–2010
2011–Present
• Classification and epidemiology of serious cutaneous ADRs
• Amoxicillin-clavulanate DILI and HLA-DRB1*15:01 association (1999)
• CBZ-induced SJS/TEN and HLA-B*15:02 in Han Chinese
• Allopurinol SJS/TEN/DIHS and HLA-B*58:01 in Han Chinese
• NVP rash–associated hepatitis and HLA-DRB1*01:01
• PREDICT-1 demonstrate that HLA-B*57:01 eliminates ABC hypersensitivity syndrome (AHS) and SHAPE demonstrates generalizability across black and white ethnicity
• Observational studies confirm utility of HLA-B*57:01 screening in clinical practice
• Routine HLA-B*57:01 screening for AHS
• HLA-A*02:01/DRB1*15:01/DQB1*06:02 and AC DILI
• HLA-A*31:01 with CBZ-induced SJS/TEN, DIHS, MPE in Caucasians and Japanese
• Standardization of severe cutaneous reactions
• Association between HLA-C*04:01 and NVP SJS/TEN
• Association between HLA-DRB1*01:02 and HLA-B*58:01 with NVP hepatitis in black South Africans

FIGURE 21.3 **Timeline of HLA associations with DHSs.** Shown are the key discoveries of HLA-associated DHRs, and their translation.

(AC), and antituberculous drugs such as isoniazid and pyrazinamide, among others [41,42]. This syndrome, termed *drug-induced liver injury* (DILI) may manifest as hepatocellular injury or cholestasis (obstruction of normal bile flow) and has been shown to progress to fulminant hepatic failure in rare instances [43]. The mechanisms underlying these pathologies may include both a toxic and an immunological basis [42,43]. Isolated pancreatitis or tubulointerstitial nephritis are additional examples of drug hypersensitivity reactions with single-organ involvement.

21.4 IMMUNOPATHOGENESIS OF DRUG HYPERSENSITIVITY SYNDROMES

The immune mechanisms underlying DHS just described have been the subject of intense research, and a multitude of data now suggests that T-cell immune responses are central to pathogenesis (Figure 21.3). The initial observations that many delayed drug-induced hypersensitivity reactions occur 2–6 weeks following first drug exposure, resolve with drug discontinuation, and recur rapidly with reexposure, support this hypothesis. This pattern of rapid symptom onset at drug rechallenge suggests that naïve T cells are primed on initial drug exposure and generate a memory pool of T cells that are restimulated on reexposure to the drug, mirroring the natural immune response to infectious pathogens.

Laboratory studies have demonstrated that circulating T cells isolated from individuals with DHS proliferate in response to specific drugs or their metabolites and that these reactions are restricted to peptide presentation in the context of MHC [44–47]. Additionally, the direct observation of cytotoxic T lymphocytes in biopsy specimens of the rash or in positive patch test from patients with cutaneous drug eruptions suggests that these cells play a role in DHS [48–51]. Preexisting cross-reactive memory T-cell responses may also play a role. For instance, in the case of abacavir, some patients who were later found to have patch test–confirmed abacavir hypersensitivity syndrome demonstrated onset of clinical symptoms within 1.5 days of first drug exposure, and 25% of patients in the PREDICT-1 study developed symptoms of hypersensitivity within 5 days of starting abacavir.

These findings raise interesting questions regarding how small-molecule pharmaceutical compounds might stimulate such immune activation. Multiple models may explain these phenomena and have been described—namely, the hapten/prohapten model, the pharmacological interaction (p-i) model, and the altered peptide repertoire model (Figure 21.4).

21.4.1 The Hapten/Prohapten Model

A hapten is a small molecule (e.g., drug or drug metabolite) that covalently binds to a larger protein, thereby altering its structure or chemical composition. The haptenated protein is then hypothesized to undergo intracellular processing to generate chemically modified peptides that are incorporated onto MHC proteins for presentation to T cells. These modified peptides are antigenically distinct from the pool of "self"-antigens and therefore elicit a T-cell response. The hapten/prohapten model also postulates that some chemically inert drugs (prohaptens) may be metabolized to reactive molecules that later serve as haptens [48,52].

Examples of hypersensitivity associated with hapten modification of endogenous proteins include the binding

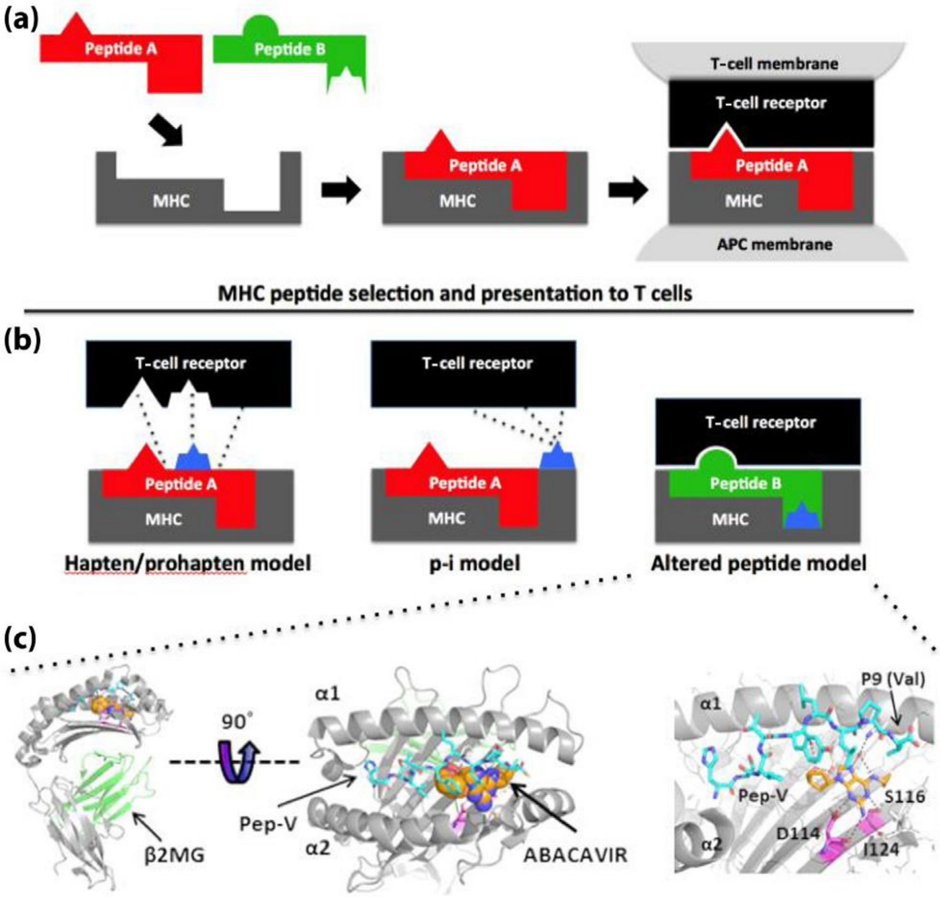


FIGURE 21.4 **Models of T cell activation by small-molecule antigens.** (a) Peptide selection and presentation by MHC. Peptide antigens are bound to the MHC protein in the intracellular environment and expressed on the surface of the antigen-presenting cell. The specificity of the MHC-peptide interaction is determined by the amino acid sequence of the MHC binding cleft and results in the selection of a distinct repertoire of bound peptides. In the example shown, peptide A is able to bind MHC but peptide B is not; thus, B is excluded from the repertoire of ligands presented in the context of this HLA allotype. The interaction of the MHC-bound peptide and the T-cell receptor, along with costimulatory signals, elicits an immune response. (b) Mechanisms of T-cell activation by small molecules. Three models have been proposed to explain T-cell stimulation by small-molecule pharmaceuticals. The hapten/prohapten model postulates that the drug binds to an intracellular protein that then undergoes antigen processing to generate haptenated peptides. These haptenated peptides are incorporated onto and presented by MHC. The haptenated-peptide is recognized as a neo-antigen to stimulate a T-cell response. The p-i model proposes that a small molecule can directly stimulate T cells in a peptide-independent fashion. The altered peptide model postulates that a small molecule can bind noncovalently to the MHC binding cleft to alter the specificity of peptide binding. This results in the presentation of novel peptide ligands to elicit an immune response. (c) The abacavir-HLA-B*57:01 interaction abides by the altered peptide repertoire model. Shown is the crystal structure of the HLA-B*57:01 (gray) oriented with the peptide binding groove facing up in panel C1 and as viewed from above in C2. Abacavir is represented as colored spheres: orange for carbon, blue for nitrogen, and red for oxygen. The peptide HSITYLLPV is shown in cyan. Hydrogen bond interactions between abacavir and HLA-B*57:01 and peptide are shown as black dashes. The amino acids that differ between HLA-B*57:01 and HLA-B*57:03, which does not participate in abacavir-associated hypersensitivity, are highlighted in magenta [55].

of penicillin derivatives to serum albumin [53] and the modification of endogenous proteins by the nitroso-sulfamethoxazole metabolite of sulfamethoxazole [54]. A time-delayed immune response is inherent to this concept, as presentation of haptenated peptides requires that the modified protein move through all steps of antigen processing, MHC loading, and traffic to the cell surface prior to T-cell interaction. It is noteworthy that examples of specific antibody responses (IgE and/or IgG) to the drug or reactive metabolite have been observed.

21.4.2 The p-i Model

The p-i model postulates that some chemically inert drugs may bind noncovalently to the TCR or MHC protein in the absence of a specific peptide ligand to directly activate T cells. This theory is supported by evidence that certain drugs are able to elicit T-cell responses following glutaraldehyde fixation of APCs, which eliminates intracellular antigen processing. It is also supported by evidence that, in some instances, T-cell responses are elicited within seconds of drug exposure, much more rapidly than would be expected if antigen processing were required [44,56,57]. The finding that inhibition of the metabolism of a drug to a chemically active molecule promotes the T-cell response in some cases suggests that inert drug compounds participate in these reactions [38]. Additionally, cell culture data showing that some drug-induced T-cell responses are abrogated when cells are washed to remove the drug implies that the chemical interactions involved are noncovalent [44,56]. The p-i model might explain DHSs that are observed at first encounter with a drug without the sensitization phase required for reactions adhering to the hapten/prohapten concept [52,59]. Examples of DHSs thought to occur via the p-i model include those for lidocaine, lamotrigine, and sulfamethoxazole.

21.4.3 The Altered Peptide Repertoire Model

The altered peptide repertoire concept postulates that small molecules may occupy sites within the peptide binding groove of MHC proteins and thereby change the chemistry and topography of the binding cleft. This, in turn, alters the specificities of MHC-peptide binding and results in selection of self-peptide antigens that are distinct from those normally bound by the unaltered MHC protein [9,10,60]. The specificity of the MHC-drug interaction is integral to this model, and this concept likely explains the numerous reports of strong associations between certain MHC alleles—often only separated from closely related alleles by micropolymorphisms in the binding cleft—with drug-specific DHS [45,47]. Further, recent work has demonstrated that this mechanism is central to the pathogenesis of abacavir hypersensitivity and carbamazepine-induced SJS/TEN in the context of HLA-B*57:01 and HLA-B*15:02, respectively [55,60,61]. These specific examples are discussed at length in the subsequent sections of this chapter.

21.5 HLA AND HYPERSENSITIVITY REACTIONS

Numerous reports in the past decade have described associations between HLA allele carriage and DHS (Table 21.2). They suggest that there exists an opportunity to predict which patients will suffer a DHS in response to drug exposure and to implement screening tests in the clinic to prevent these reactions. Of the reported associations, those involving abacavir, carbamazepine, allopurinol, nevirapine, and AC have been best characterized.

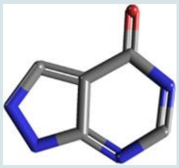
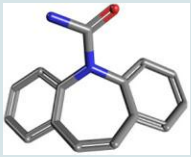
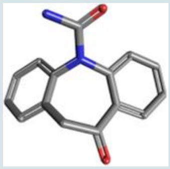
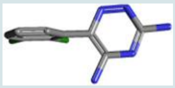
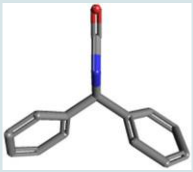
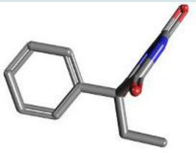
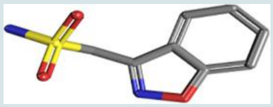
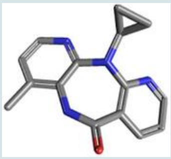
21.5.1 Abacavir

Abacavir is a nucleoside reverse-transcriptase inhibitor that is used in combination with other antiretroviral agents for the treatment of HIV-1 infection. In initial studies it was described that 5–8% of patients developed an AHR within a day and out to 6 weeks following abacavir initiation. AHR was characterized by fever, malaise, gastrointestinal and/or respiratory symptoms [62]. Rash, which is a common and early manifestation of many DHS, is absent in up to 30% of patients with abacavir hypersensitivity and often occurs later in the course of disease [63]. This nonspecific constellation of symptoms mimics the features of many common infections, reactions to other drugs, or a general inflammatory state, as seen with immune restoration following initiation of HIV-1 treatment. It therefore may pose a diagnostic challenge to the clinician. Immediate discontinuation of abacavir is mandatory, and reexposure elicits rapid return of symptoms, which are often more severe than the initial episode and potentially life threatening.

In 2001, it was observed that abacavir hypersensitivity reactions (AHRs) occur more frequently in persons of Caucasian ancestry as opposed to black or Asian individuals [64]. Subsequent studies demonstrated an association between AHR and the MHC class I allele, HLA-B*57:01 [65,66]. The original study involved a recombinant mapping strategy to refine the association to HLA-B*57:01, given the strong LD with HLA-C*06:02 and HLA-DRB1*07:01 observed in Caucasian populations.

HLA-B*57:01 is most prevalent in persons of European descent (5–8% carriage rate) and less frequently found in persons of African American descent (2.4% carriage rate) or other ethnic groups [67] (Figure 21.5). This is consistent with the hypothesis that HLA-B*57:01 expression is linked to AHR. The immunologic basis of AHR was proven through the use of epicutaneous patch-testing studies that demonstrated a marked response to abacavir when applied to the skin of hypersensitive individuals [51,68,69]. Additionally, *ex vivo* studies demonstrated that CD8+ T cells derived from abacavir-hypersensitive patients proliferate and secrete proinflammatory cytokines in response to abacavir stimulation of HLA-B*57:01 expressing APCs, which further supports the concept that MHC-T cell interactions are mechanistic for AHR [70,71]. Seminal studies include the PREDICT-1 trial, which randomized nearly 2000 HIV-positive patients, most of whom were Caucasian, to undergo either prospective HLA-B*57:01 screening with subsequent exclusion of abacavir therapy for HLA-B*57:01 positive patients or abacavir initiation with close clinical monitoring and retrospective HLA-B*57:01 genotyping for patch test–confirmed AHR cases. Pretreatment HLA-B*57:01 screening eliminated patch test–confirmed AHR (0% AHR incidence in prospective screening group vs. 2.7% AHR incidence in controls) with a 100% negative predictive value [68]. The SHAPE trial followed soon thereafter and showed that HLA-B*57:01 screening was 100% sensitive as a marker for immunologically confirmed AHR in black patients [69]. In both studies, the positive predictive value of HLA-B*57:01 carriage for the development of AHR was ~50%, suggesting that

TABLE 21.2 HLA-Associated Drug Hypersensitivity Reactions

Drug	Structure	Alleles	Year Described	Populations
SJS/TEN:				
Allopurinol		B*58:01 (or B*58 haplotype) [72-77]	1989	Han Chinese, Thai, European, Italian, Korean
Carbamazepine		B*15:02 [78-88]	2004	Han Chinese, Thai, Malaysian, Indian
		B*15:11 [89,90]	2010, 2011	Korean, Japanese
		B*15:18, B*59:01 and C*07:04 [91]	2010	Japanese
		A*31:01 [89,92-94]	2011	Japanese, Northern European, Korean
Oxcarbazepine		B*15:02 and B*15:18 [95,96]	2010	Han Chinese, Taiwanese
Lamotrigine		B*15:02 (positive) [95]	2010	Han Chinese
		B*15:02 (no association found) [97,98]		Han Chinese
Phenytoin		B*15:02, B*13:01, Cw*08:01 and DRB1*16:02 [81,87,95]	2007, 2010	Han Chinese
Phenobarbital		B*51:01 [99]	2013	Japanese
Zonisamide		B*51:01 [99]	2013	Japanese
Nevirapine		HLA-C*04:01 [100]	2013	Malawian

(Continued)

TABLE 21.2 Continued

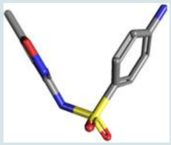
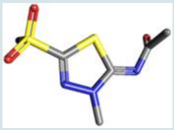
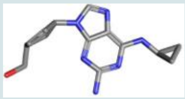
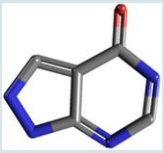
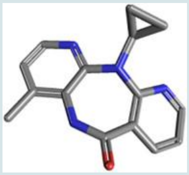
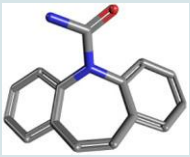
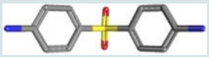
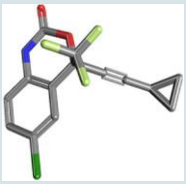
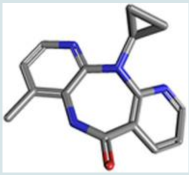
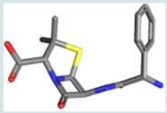
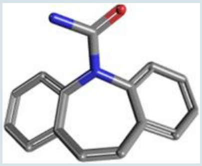
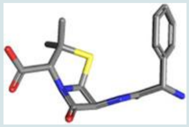
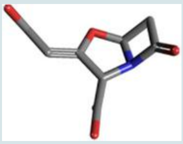
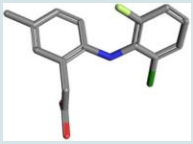
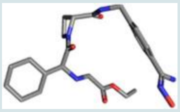
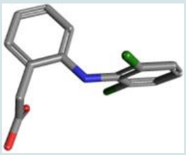
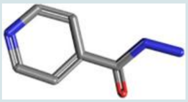
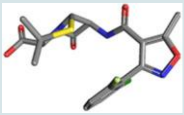
Drug	Structure	Alleles	Year Described	Populations
Sulfamethoxazole		B*38 [99]	2008	European
Methazolamide		B*59:01, Cw*01:02 alleles and B*59:01-Cw*01:02-A*24:02 haplotype [102]	2010	Korean and Japanese
Sulphonomides (sulfa antimicrobials)		A*29, B*12 and DR7 [103]	1987	European
Oxicam		B*73, A*2 and B*12 [101,103]	1987, 2008	European
HSS/DIHS/DRESS:				
Abacavir		B*57:01 [65,66,68,69]	2002	European, African
Allopurinol		B*58:01 (or B*58 haplotype) [73-77,104-106]	1989, 2005	Han Chinese, Korean, Japanese, Thai and European
Nevirapine		DRB1*01:01(hepatitis and CD4%≥25) and DRB1*01:02 (hepatitis) [107,108]	2005 2013	Australian, European and South African
		Cw*8 or Cw*8-B*14 haplotype [109,110]	2006, 2007	Italian and Japanese
		Cw*4 (rash) [108,111]	2011, 2012	Blacks, Asians, Whites Han Chinese
		B*35 [108], B*35:05 (rash) [112]	2011	Asian
Carbamazepine		8.1 AH (HLA A*01:01, Cw*07:01, B*08:01, DRB1*03:01, DQA1*05:01, DQB1*02:01) [113]	2006	Caucasians
		A*3101 [89,92-94]	2011	Northern European, Japanese and Korean
		A*11 and B*51 (weak) [93]	2012	Japanese
Dapsone		B*13:01 [114]	2013	Han Chinese
Delayed Rash:				
Efavirenz		DRB1*01 [115]	2008	French

TABLE 21.2 Continued

Drug	Structure	Alleles	Year Described	Populations
Nevirapine		DRB1*01 [115]	2008	French
		Cw*04 [108,116]	2009, 2011	African, Asian, European and Thai
		B*35:05, rs1576*G CCHCR1 status [112,117]	2009	Thai
Aminopenicillins	 amoxicillin	A*2 and DR*52 [118]	1998	Italian
Carbamazepine (including MPE)		A*31:01 [85,92]	2006	Han Chinese and Northern European
DILI				
Amoxicillin-clavulanate	 Amoxicillin	 Clavulanate	1999, 2000, 2011	European
	Alleles: DRB1*15:01, DRB107 protective, HLA-A*02:01, HLA-DQB1*06:02 and rs3135388, a tag SNP of HLA-DRB1*15:01—DQB1*06:02 [119-121]			
Lumiracoxib		HLA-DRB1*15:01—HLA-DQB1*06:02—HLA-DRB5*01:01—HLA-DQA1*01:02 haplotype [122]	2010	International, multicenter
Ximelagatran		DRB1(*)07 and DQA1(*)02 [123]	2008	Swedish
Diclofenac		abacavirB11, C-24T, UGT2B7*2, IL-4 C-590-A [42,43,124]	2007	European
Isoniazid		NAT2 slow acetylator, CYP2E1*5 and *1B [42,43]	2000, 2003, 2006	European
Flucloxacillin		B*57:01, DRB1*01:07—DQB1*01:03 [42,125]	2009	European

(*Continued*)

TABLE 21.2 Continued

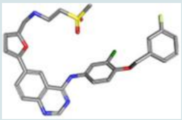
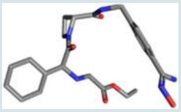
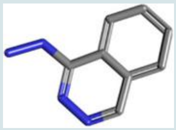
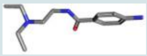
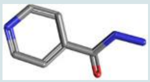
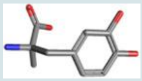
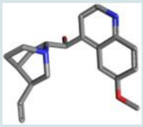
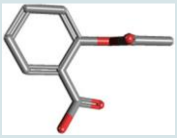
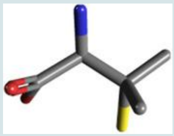
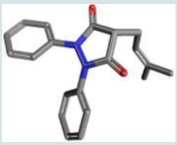
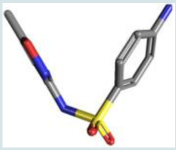
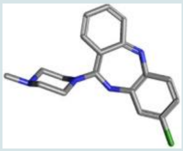
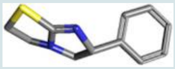
Drug	Structure	Alleles	Year Described	Populations
Lapatinib		DRB1*07:01—DQA2*02:01—DQB1*02:02/02:02 [126]	2011	International, multicenter
Ximelagatran		DRB1*07 and DQA1*02 [123]	2008	European
Other Drug Reactions:				
Hydralazine		Drug-induced lupus erythematosus: DR*4 [127]	1980	European
Procainamide, Isoniazid, Methyldopa and Quinine	 Procainamide  Isoniazid  Methyldopa  Quinine	Drug-induced lupus erythematosis: DR*4 [128]		Italian
Aspirin		Urticaria/angioedema: DRB1*13:02—DQB1*06:09—DPB1*02:01 haplotype [129]	2005	Korean
		Asthma: DPB1*03:01 [130]	2004	Korean
Gold sodium thiomalate		Mucocutaneous reaction: DR*5 [131]	1994	Spanish
		Proteinuria, thrombocytopenia or leukopenia: B*8 or DR*3 [132]	1985	European
NSAIDs		Anaphylactoid and cutaneous reactions: DR*11 [133]	1999	Spanish
D-penicillamine		Myasthenia gravis: DR*1 [134]	1983	Mixed Caucasian
Beryllium		Granulomatous lung disease: HLA-DPB1 gene and DPR1 gene polymorphisms, DRB1*13 and DQB1*06 [135,136]	2000, 2003	North American
Feprazone		Fixed drug eruption: B*22 [137,138]	1994	Italian

TABLE 21.2 Continued

Drug	Structure	Alleles	Year Described	Populations
Sulfamethoxazole		Fixed drug eruption: A*30—B*13—Cw*6 haplotype [139]	2001	Turkish
Clozapine		Agranulocytosis: B*38, DR*4 and DQw*3 [140,141]	1990, 1998	Jewish
		(6672G>C) in HLA-DQB1 [142]	2011	North American
		Cw71B*18 or B*39 or B*44/DRB*5 [143]	2007	Caucasian
Levamisole		Agranulocytosis: B*27 [144]	1990	South American

Source: Adapted from Pavlos et al. [10].

Ig—immunoglobulin; FcR—Fc receptor; NK cells—natural killer cells; DRESS—drug reaction with eosinophilia and systemic symptoms; DIHS—drug-induced hypersensitivity syndrome; HSS—hypersensitivity syndrome; SJS—Stevens-Johnson syndrome; TEN—toxic epidermal necrolysis; DILI—drug-induced liver injury; AGEP—acute generalized exanthematous pustulosis. Chemical Structures: blue: nitrogen; red: oxygen; yellow: sulphur; grey: hydrogen; black: carbon.

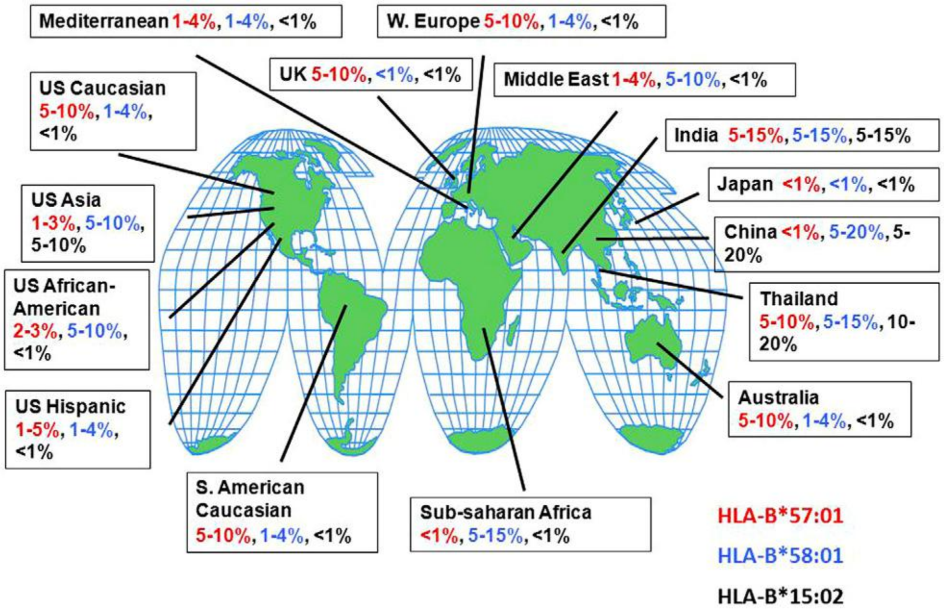

Mediterranean 1-4%, 1-4%, <1%
W. Europe 5-10%, 1-4%, <1%
US Caucasian 5-10%, 1-4%, <1%
UK 5-10%, <1%, <1%
Middle East 1-4%, 5-10%, <1%
India 5-15%, 5-15%, 5-15%
US Asia 1-3%, 5-10%, 5-10%
Japan <1%, <1%, <1%
China <1%, 5-20%, 5-20%
US African-American 2-3%, 5-10%, <1%
Thailand 5-10%, 5-15%, 10-20%
US Hispanic 1-5%, 1-4%, <1%
Australia 5-10%, 1-4%, <1%
S. American Caucasian 5-10%, 1-4%, <1%
Sub-saharan Africa <1%, 5-15%, <1%
HLA-B*57:01
HLA-B*58:01
HLA-B*15:02

FIGURE 21.5 **Frequency of HLA alleles across diverse populations.** HLA carriage frequency for key HLA-B alleles associated with drug hypersensitivity show marked variation across different ethnic populations. HLA carriage frequencies for HLA-B*57:01 were obtained from Nolan et al. [145]. Carriage frequencies for HLA-B*58:01 and HLA-B*15:02 were calculated from allele frequency data obtained from www.allelefrequencies.net. Note that within-country or within-region data are not shown and the final range indicated is inclusive of most studies or is taken from datasets containing >1000 subjects. *Source: adapted from Phillips et al. [201].*

the HLA-B*57:01 allotype is necessary but not sufficient for AHR [68,69]. It is not clear at present why some HLA-B*57:01 carriers are abacavir-tolerant, and this question is currently under investigation. Through these later studies that utilized patch testing to identify true immunologically-mediated AHR it was also determined that the window for presentation of true-immunologically mediated AHR was within 1.5 days to 3 weeks following abacavir initiation.

The mechanisms underlying AHR have recently been illuminated and are consistent with an altered peptide repertoire model in which abacavir undergoes metabolism-independent, direct and noncovalent interaction with the HLA-B*57:01 binding cleft to allow presentation of novel endogenous peptides (Figure 21.4(c)). The best evidence to support this hypothesis comes from two structural analyses of the abacavir-peptide-HLA-B*57:01 trimolecular complex. In each analysis, unmodified abacavir was shown to adopt an extended conformation within the HLA-B*57:01 peptide binding groove and extensive, noncovalent contacts with the HLA molecule; these contacts include residues within multiple binding pockets (C, D, E, and F pockets). This effectively abrogates peptide binding at these sites, although peptides were shown to maintain contact at the preferential anchor site in the B pocket of the binding cleft. These structures also reveal that bound peptides make very few contacts with the abacavir molecule itself [55,61]. In essence, abacavir association with the HLA-B*57:01 molecule changes the chemistry and topography of the peptide binding groove and results in presentation of an altered repertoire of T-cell epitopes to which the individual has not been tolerized. The "altered self" repertoire of peptides elicits a polyclonal T-cell response that clinically manifests as AHR [55].

The model just described explains many previous findings associated with AHR. First, the abacavir-HLA-B*57:01 molecular contacts identified in the crystal structures are consistent with mutagenesis data demonstrating that certain residues in the peptide binding cleft—namely Asp114 and Ser116—are required to elicit T-cell responses to the abacavir-peptide-HLA-B*57:01 trimolecular complex [45]. Second, elucidation of these contact sites provides an explanation for the absence of AHR seen in individuals with the closely related allotypes HLA-B*58:01, HLA-B*57:02, and B*57:03. In each of these cases, polymorphisms at key residues in the binding cleft prevent abacavir association, and thus peptide selection remains unchanged in the presence of the drug. Third, it has been shown that the peptides eluted from abacavir-treated APCs expressing HLA-B*57:01 in cell culture models display broad sequence diversity and include many neo-epitopes when compared to peptides eluted from abacavir-naïve cells [55,60,61]. Illing et al. demonstrated that 20–25% of peptides recovered from abacavir-treated HLA-B*57:01 APCs were distinct from those isolated from untreated cells, illustrating the dramatic shift in the repertoire of HLA-B*57:01-bound peptides in the presence of abacavir [61]. Further, analysis of the specificities of the T-cell responses elicited by abacavir-treated APCs expressing HLA-B*57:01 demonstrates polyclonal TCR usage consistent with broad diversity in the pool of peptides presented in this context [9,45,61,71]. The simultaneous requirement that the involved peptides be of high sequence diversity and be HLA-B*57:01-restricted is not reconcilable by the hapten/prohapten concept. HLA-B*57:01, B*58:01, B*57:02, and B*57:03 have extensive overlap in their binding specificities; thus, if hapten modification of peptide were central to AHR pathogenesis, we would expect to isolate one or only a few unique HLA-B*57:01-restricted peptides in elution experiments and observe a narrow range of TCR usage in response to stimulation by the abacavir-peptide-HLA-B*57:01 trimolecular complex. Additionally, no abacavir-haptenated peptides were identified in any study, as would be expected if AHR adhered to the hapten/prohapten concept. Finally, HLA-B*58:01, B*57:02, and B*57:03 alleles are found at relatively high frequencies among blacks, a population in which HLA-B*57:01 carriage and AHR are less frequently observed [61,147] (Figure 21.5). Thus, the epidemiologic data demonstrating that AHR is restricted to the HLA-B*57:01 allotype is explained by the specific association of abacavir with the HLA-B*57:01 molecule, as illustrated by structural studies.

21.5.2 Carbamazepine

Carbamazepine is an aromatic amine anticonvulsant used to treat epilepsy, bipolar disorder, and trigeminal neuralgia. It has been associated with multiple types of DHS, including SJS/TEN, maculopapular exanthema, and DIHS. Carbamazepine-associated ADRs typically manifest approximately 2–8 weeks after initiation of the drug and can include rash, fever, lymphadenopathy, and internal organ involvement [148]. These reactions are estimated to occur in 1–6 of every 10,000 individuals receiving this drug in European populations [92].

However, the prevalence of carbamazepine-associated ADRs in persons of Southeast Asian descent is approximately ten-fold higher than in Europeans [10,78]. In 2004, Chung et al. demonstrated the first association of carbamazepine-induced ADRs with a specific HLA genotype. Remarkably, their study revealed that the HLA-B*15:02 allele was present in 100% of Han Chinese patients with carbamazepine-associated SJS/TEN, as compared to only 3% of carbamazepine-tolerant controls and 8.6% of the general population. Based on these data, the negative predictive value of HLA-B*15:02 for carbamazepine-associated SJS/TEN approximates 100% [78]. Multiple subsequent studies have confirmed the association of the HLA-B*15:02 allotype and carbamazepine-induced SJS/TEN in Chinese, Thai, Malaysian, and Indian populations [79-84,149] (Table 21.2).

In 2009, Chen et al. demonstrated that prospective screening for HLA-B*15:02 positivity among 4855 Han

Chinese patients and subsequent avoidance of carbamazepine in these individuals eliminated cases of carbamazepine-induced SJS/TEN in this group. Based on historical incidence rates, the authors estimated that ten cases of carbamazepine-associated SJS/TEN would have (0.25%) developed in the absence of HLA-B*15:02 screening; thus their report supports a role for pretreatment HLA-B*15:02 genotyping prior to initiation of carbamazepine [150].

The observation that not all HLA-B*15:02 carriers develop hypersensitivity in response to carbamazepine exposure prompted Ko et al. to characterize the TCR repertoire derived from patients with carbamazepine-associated SJS/TEN. These studies demonstrated that a dominant TCR clonotype, VB-11-ISGSY, was present in clinical samples from 84% of HLA-B*15:02–positive patients with active carbamazepine-induced SJS/TEN and 14% of healthy, carbamazepine-naïve controls. Interestingly, this clonotype was absent in all carbamazepine-tolerant controls, including two HLA-B*15:02 carriers. Cell culture experiments demonstrated that T cells derived from individuals who were carbamazepine naive but positive for HLA-B*15:02 and VB-11-ISGSY acquired a cytotoxic phenotype following stimulation with carbamazepine. It was also demonstrated that these responses were blocked with the addition of a VB-11-ISGSY–specific antibody. These studies imply that, in the case of HLA-B*15:02-restricted carbamazepine-SJS/TEN, the presence of both the HLA-B*15:02 allotype and a specific TCR clonotype is predictive of the development of carbamazepine-induced ADR [151]. There are likely other, as yet unknown, contributing factors given that some individuals with a history of carbamazepine SJS/TEN in this study were found to be negative for both HLA-B*15:02 and the VB-11-ISGSY TCR clonotype.

HLA-B*15:02 is carried in <0.1% of the Caucasian population (Figure 21.5). More recently, carriage of the HLA-A*31:01 allele has also been identified as predictive of carbamazepine-induced SCAR in persons of Northern European ancestry [92] (Table 21.2). In a GWAS of 145 patients with carbamazepine-induced hypersensitivity (including SJS/TEN and other DIHSs) and 257 control patients, McCormick et al. demonstrated that the presence of the HLA-A*31:01 allele predicted the development of carbamazepine-induced SCAR with an odds ratio of 9.12 ($p = 1.0 \times 10^{-7}$; test sensitivity 26% and specificity 96%) [92]. This association was also demonstrated in Japanese and Korean patients [93,94,152]. In the Japanese population, it is estimated that HLA-A*31:01 carriage is 60.7% sensitive and 87.5% specific as a predictor of carbamazepine-induced SCAR [94]. It is estimated that the prevalence of HLA-A*31:01 carriage is 2–5% in Northern Europeans, 7% in Han Chinese, 7–12% in Japanese populations, and 5% in persons of Korean ancestry (see www.allelefrequencies.net).

In 2012, Yip et al. published a meta-analysis of 23 studies describing associations between carbamazepine-related cutaneous ADRs and HLA genotypes across multiple populations [153]. The aggregate data included HLA allotype information on 690 patients with carbamazepine-related cutaneous ADRs, 1585 carbamazepine-tolerant controls, and 7056 population controls. Their findings demonstrate that, in persons of Asian descent, carbamazepine-SJS/TEN and HLA-B*15:02 carriage are associated with a pooled odds ratio (OR) of 113.4 (95% $\text{CI} = 51.2 - 251.0$, $p < 1 \times 10^{-5}$). These results imply that HLA-B*15:02 positivity is 96% sensitive and 88% specific as a predictor of carbamazepine-induced SJS/TEN in this population. The calculated positive and negative predictive values were 1.8% and 100%, respectively, and it is therefore estimated that the number needed to screen to prevent one case of carbamazepine-associated SJS/TEN is 461 in persons of Asian ancestry.

Unlike the association of HLA-A*31:01 with multiple phenotypes of carbamazepine-DHS in Caucasians, there has not been a convincing association between carbamazepine-induced HSS or MPE and HLA-B*15:02 carriage in Caucasians and this is likely driven by the fact that the prevalence of HLA-B*15:02 is rare in this population. Yip et al. also confirmed the association between HLA-A*31:01 positivity and carbamazepine-related cutaneous ADRs with a pooled OR of 9.5 (95% $\text{CI} = 16.1 - 13.9$, $p < 1 \times 10^{-5}$) [153]. In subgroup analyses, it was shown that HLA-A*31:01 carriage was 58% sensitive and 87% specific as a predictor of carbamazepine-associated ADRs in the Japanese population, and 26% sensitive and 96% specific in Caucasians.

With regard to the molecular mechanisms underlying carbamazepine-associated ADRs, *in vitro* studies have shown that carbamazepine presentation to T cells in the context of the HLA-B*15:02 protein is independent of intracellular drug or antigen processing, but does require peptide binding to stabilize the class I MHC complex on the cell surface [47]. Interestingly, similar experiments involving APCs that express other members of the HLA-B75 family, which includes HLA-B*15:02, demonstrated T-cell responses in the context of carbamazepine presentation by these related allotypes. This finding suggests that the HLA-B75 serotypes may share a common conformation or TCR that allows for carbamazepine binding [47]. Consistent with this laboratory data, carbamazepine-induced SJS/TEN has been observed in association with HLA-B75 serotypes B*15:08, B*15:11, and B*15:21 in Indian, Japanese, and Korean cohorts [89,90,149] (Table 21.2). The region of carbamazepine association with HLA-B*15:02 maps to the B pocket of the peptide binding groove with a likely primary contact at the Arg62 residue—a conserved amino acid among B75 serotypes. Additional contacts at the Asn63, Ile95, and Leu156 residues also likely participate in carbamazepine-B*15:02 interactions, as mutagenesis studies demonstrate reduced affinity carbamazepine binding in their absence [47]. These findings are consistent with a direct, non-covalent and metabolism-independent mechanism of carbamazepine association with the HLA-B*15:02 molecule.

Analysis of peptides eluted from carbamazepine-treated APCs expressing HLA-B*15:02 demonstrated that 15% of isolated peptides were distinct from those obtained in untreated cells [61]. This finding is similar to that of the abacavir-B*57:01-peptide interaction and suggests that the altered peptide repertoire model may also be etiologic for carbamazepine-induced SJS/TEN. Additional work is needed to precisely define the nature of carbamazepine-HLA-B*15:02–peptide interactions.

Multiple questions remain regarding the genetic mechanisms that contribute to carbamazepine-induced ADRs. For example, a growing body of data suggests that the ADRs associated with carbamazepine exposure in the context of HLA-A*31:01 are phenotypically distinct from those elicited in the context of HLA-B*15:02. HLA-A*31:01 positivity appears to be associated more frequently with DRESS/DIHS/HSS and maculopapular exanthema (MPE), whereas HLA-B*15:02 carriage is associated with more severe cutaneous reactions such as SJS/TEN in Asian populations [92,154]. SJS/TEN, DIHS/DRESS/DHS, and MPE are distinct syndromes, so this may not be surprising; however, the mechanisms underlying this discrepancy are not clear at present. Study design may be important, and, in particular, discrepancies may result from studies that have used GWAS or SNP approaches to identify HLA associations without full high-resolution HLA A B C and DR/DQ typing. These methodologies may miss important associations, particularly those due to the non-specificity of LD within HLA. Further, numerous reports have demonstrated examples of carbamazepine-induced SJS/TEN in HLA-B*15:02–negative patients, which suggests that other HLA class I and particularly HLA-B alleles may be important [153,155,156]. This indicates that additional factors that may be associated with this disease are yet to be identified.

21.5.3 Allopurinol

Allopurinol is a xanthine oxidase inhibitor that is used to treat hyperuricemia and is associated with DIHS and SJS/TEN in some patients. The first report of a potential genetic association between allopurinol administration and observed ADRs occurred in 1989 when it was shown that the AW33 and BW58 loci were prevalent in Southern Chinese patients with cutaneous manifestations of allopurinol hypersensitivity [72]. This observation was followed many years later by a case-controlled GWAS that identified the HLA-B*58:01 genotype as strongly associated with allopurinol-induced SJS/TEN in the Han Chinese population. In this report, the HLA-B*5801 allotype was present in 100% of cases of allopurinol-SJS/TEN/DRESS/DIHS/HSS, 15% of allopurinol-tolerant patients, and 20% of healthy controls [73]. These findings have since been replicated in Japanese, Thai, and Korean populations [74,75,104,157,158] (Table 21.2). However, the association between HLA-B*58:01 carriage and allopurinol-SJS/TEN was found to be less robust in Caucasian populations, a group in which the prevalence of the HLA-B*58:01 allele is much lower (1-6% carriage rate in Caucasians as compared to 15% in Han Chinese) (allele-frequencies.net and Figure 21.5). In a European cohort, only 61% of patients with allopurinol-induced SJS/TEN carried the HLA-B*58:01 gene, suggesting that additional genetic factors might contribute to the development of allopurinol ADRs in this group [101].

In 2011, Genin et al. conducted a GWAS using >500 sera collected from patients enrolled in the RegiSCAR database, a European registry of samples collected from patients with SCARs across the SJS/TEN spectrum. This group included 57 patients with SJS/TEN following allopurinol exposure. A study by Genin et al. identified six SNPs that were significantly associated with allopurinol-induced SJS/TEN, all of which were in LD with the HLA-B loci [76]. Similarly, a 2011 GWAS analysis of Japanese patients with allopurinol-SJS/TEN identified multiple SNPs with strong association to this ADR, some of which were in complete LD with HLA-B*58:01 [105]. Interestingly, both studies identified SNPs found in high prevalence in patients with allopurinol-induced ADRs that overlap genes known to be associated with autoimmune inflammatory conditions (psoriasis, rheumatoid arthritis) and immune function [105,159-161]. Importantly, studies of allopurinol-related ADRs in these European populations show that the association with HLA-B*58:01 is significant but incomplete, which may be due to the lower frequency of HLA-B*58:01 carriage in these groups and/or additional contributing factors that have not been identified [77].

The mechanisms of allopurinol-induced ADR in the context of HLA-B*58:01 expression have not yet been elucidated, and functional studies have been hard to replicate. Unlike carbamazepine and nevirapine, where clinical phenotype and ethnicity seem to be important drivers of the HLA association, HLA-B*58:01 has been associated with both the SJS/TEN and DIHS/DRESS/HSS phenotype across different populations. It is predicted that, as with abacavir and carbamazepine, residues in the antigen binding cleft of the HLA molecule are involved in the HLA-drug interactions. Similar to the exquisite specificity of HLA-B*57:01 for abacavir hypersensitivity and tolerance in patients with closely related B17 serotype alleles such as HLA-B*57:03, B*57-02, B*58:01, HLA-B*58:02, there is a closely related allele that is not associated with allopurinol hypersensitivity and differs from the HLA-B*58:01 allele at six amino acids, three of which are located in the peptide binding groove. Although the F pocket of the peptide binding groove is predicted to be compatible with allopurinol binding, allopurinol is a smaller molecule than the other drugs involved in

DHS. More recent functional studies suggest that it may be allopurinol's metabolite oxypurinol which mediates DHS however the exact mechanisms of this interaction remain to be defined [162].

21.5.4 Nevirapine

Nevirapine (NVP) is a nonnucleoside reverse-transcriptase inhibitor that is part of combination antiretroviral therapy for the treatment of HIV-1 infection. Approximately 5% of patients will develop DHS within the first six weeks following initiation of NVP therapy. Reactions often consist of rash of varying severity (including SJS/TEN), fever and/or hepatitis, and, in rare cases, they can be fatal [163,164]. Observational studies have noted multiple host and environmental factors that influence the likelihood and severity of NVP-induced ADRs. Risk factors for severe disease that have been shown in some include prolonged exposure to any antiretroviral therapy, co-infection with hepatitis B or C viruses, and abnormal serum hepatic transaminase levels at baseline. Interestingly, lower CD4+ T-cell counts prior to initiation of NVP have been shown to be protective against the development of NVP-associated ADRs; however, more recent studies have suggested that this effect may be less important when switching to nevirapine in patients who are virologically suppressed [165-171].

Deciphering the genetic variants that influence risk of NVP-associated ADRs is not straightforward, and the cumulative data indicate that multiple genes associated with immune function and drug metabolism predispose to specific manifestations of NVP hypersensitivity in specific populations. In 2005, it was shown that the combination of higher baseline CD4% (CD4% T cells $\geq$25% of total) and the presence of the HLA-DRB1*01:01 allele is significantly associated with NVP-rash associated hepatitis in an Australian cohort of HIV-infected patients, suggesting that lower CD4% may significantly abrogate specific phenotypes of NVP-DHS [107]. Subsequent studies also demonstrated an association between HLA-DRB1*01:01 carriage and NVP-related hepatic and cutaneous reactions in other Caucasian populations [108,115]. Conversely, the class I MHC allele HLA-Cw8 was identified in Sardinian and Japanese populations in association with NVP-induced hypersensitivity, and the HLA-B*35:05 and HLA-Cw*04 variants were shown to be involved in isolated cutaneous NVP reactions in African Americans, Thais, and Han Chinese [108-112,116,117].

In 2011, Yuan et al. published the results of a large retrospective case-controlled study in which they investigated the genetic variants associated with NVP-related hypersensitivity in 276 cases from 11 countries worldwide. Although their study reported only 2-digit HLA typing, their findings support a relationship between carriage of certain class I MHC alleles—namely, HLA-B*35 and HLA-Cw*04—and the development of cutaneous NVP-related hypersensitivity in Asian populations (Table 21.2). They also found that, among whites, HLA-DRB1*01 positivity is associated with increased prevalence of hepatic and/or systemic NVP-induced DHS but not with isolated cutaneous reactions. There was no association with NVP-DHS and HLA-DRB1*01 carriage among blacks or Asians, two groups in which this allele is infrequently present. However, a subsequent South African case-control study that included high-resolution HLA ABC and DR typing in patients who developed grade 3/4 hepatitis within three months of starting NVP versus those tolerating NVP for at least three months showed an association with HLA-DRB1*01:02, an HLA class II MHC allele with a high degree of structural similarity to HLA-DRB1*01:01 [172].

More recently, a new association was reported between HLA-C*04:01 and NVP-SJS/TEN ($n = 33$) in a case-control study of NVP-treated HIV patients in a Malawi cohort consisting of 117 patients with NVP-DHS and 151 age-, sex-, and gender-matched controls [100]. Interestingly, polymorphisms in CYP2B6, the main cytochrome p450 isoform involved in NVP metabolism that shows genetic variation, was shown to be associated with increased risk of NVP-induced cutaneous ADRs in blacks and whites [108]. Additionally, in a GWAS analysis of 72 Thai patients with NVP-related ADR, it was observed that two SNPs located within the CCHCR1 gene, which regulates keratinocyte proliferation and is known to be associated with psoriasis, are predictive of NVP-induced rash [117].

These findings are similar to those showing a relationship between CCHCR1 and allopurinol hypersensitivity as discussed in the previous section. Because of the high LD between many of these SNPs and HLA-B and C alleles, they should be interpreted with caution because the likely true association is HLA-related. Without mapping the region at risk, it is possible that these associations may be missed. This is similar to the story of the abacavir and HLA-B*57:01 association, where a haplotypic marker in high LD with HLA-B*57:01, HCP5 (rs2395029), is present in most but not all HLA-B*57:01 patients with patch test–confirmed abacavir hypersensitivity; the presence of this tag SNP may vary across the population involved [173]. Taken together, these data indicate that NVP-DHS has complex associations that appear to be specific to phenotype and ethnicity and may result from a complex interplay of immunologic, genetic, and metabolic factors.

21.5.5 Amoxicillin-Clavulanate

The co-formulation of amoxicillin, an aminopenicillin antibiotic, and the beta-lactamase inhibitor clavulanic acid is one of the most frequently prescribed antimicrobials for the treatment of bacterial infections. While the overall side-effect profile of amoxicillin-clavulanate (AC) is favorable, this drug has been identified as a common cause of drug-induced liver injury (DILI) with highly variable clinical manifestations [42,174-176]. Hepatic injury may occur in a cholestatic, hepatocellular, or mixed pattern, and the onset of symptoms has been reported to range between 3 and 59 days following initial drug exposure [175,177]. Rarely, AC-DILI may progress to fulminant hepatic failure requiring liver transplantation or resulting in death. Age and duration of therapy have also been proposed to influence the phenotype of AC-DILI, with advanced age and longer treatment courses favoring the development of cholestatic or mixed patterns of hepatotoxicity [175,178].

Data suggest that both HLA class I and class II genetic variants contribute to AC-induced DILI. A Belgian study published in 1999 reported that the HLA class II haplotype DRB1*15:01-DRB5*01:01-DQB1*06:02 was present in 57% of 35 patients with cholestatic AC-DILI as compared to only 15% in population controls ($p<0.0002$) [119]. In 2000, a similar study performed in a small cohort of Scottish patients confirmed the association of cholestatic AC-DILI and HLA DRB1*15:01. DRB1*15:01 was present in 70% of 20 case patients versus 20% of controls, and it was shown to confer a 9.25 relative risk for the development AC-DILI in this analysis [120]. In concordance with these findings, a UK study published by Donaldson et al. reported that 53% of AC-induced DILI patients carried HLA DRB1*15 alleles compared to 33% of AC-tolerant patients and 30% of population controls ($p=0.002$). Also reported was a potential protective role for the HLA DRB1*07 family of alleles, as these genes were observed at reduced frequencies in AC-DILI patients compared to controls.

These studies were followed in 2011 by a GWAS analysis that included 201 European white, US, and Spanish patients. It was demonstrated that AC-DILI is strongly associated with the SNP rs9274407 that is correlated with the HLA haplotype DRB1*15:01-DQB1*06:02. Additionally, this analysis revealed an independent association with the SNP rs2523822, a known tag for the class I HLA allele A*02:01 among Northwestern Europeans. High-resolution HLA genotyping confirmed these HLA associations in US and Northwestern European populations, but failed to show an association between AC-DILI and HLA-A*02:01 in persons of Spanish descent. These findings likely reflect LD patterns, as it is known that rs2523822 and HLA-A*02:01 exist in strong LD among Northwestern European populations but not among Southern Europeans. Moreover, the associations and haplotypes may also be dependent on ethnicity and subphenotypes [121,178].

Epidemiologic data suggests that the clinical features of amoxicillin-clavulanate-associated DILI may correlate with ethnic background. For example, Lucena et al. demonstrated that 36% of Spanish AC-DILI cases are characterized by hepatocellular injury, 31% by cholestatic damage, and 33% by a mixed hepatocellular/cholestatic pattern [121,175]. This is in contrast to French and Belgian AC-DILI cases, in which a cholestatic pattern of liver injury predominates in 66–74% of cases and hepatocellular damage occurs in only 11–13% of cases [121,175,178]. Stephens et al. postulated that differential HLA associations might underlie this phenotypic variation. To investigate this possibility, they performed class I and class II HLA genotyping for 75 Spanish AC-DILI cases of which 37% were characterized by a hepatocellular toxicity pattern. In this analysis, HLA-A*30:02 and HLA-B*18:01 were associated with hepatocellular liver injury (OR 6.7 and 2.9, respectively), although the latter association did not reach statistical significance. Conversely, carriage of the class II haplotype DRB1*15:01-DQB1*06:02 was significantly associated with cholestatic or mixed AC-DILI in this group with an odds ratio of 3.0 ($p=0.014$). Interestingly, Stephens et al. note that HLA-A*30:02 and HLA-B*18:01 are both present in the MHC haplotype HLA-A*30:02-B*18:01-Cw-05:01-DRB1*03:01-DQB1*02:01, which is the third most common MHC haplotype among Southern Europeans and is less frequently found in Northern European populations [178]. Taken together, these findings suggest that different HLA class I and class II risk alleles and HLA class I/II pairings are associated with phenotypically distinct presentations of AC-DILI, and that the frequencies of allelic carriage among ethnically distinct populations determine the clinical features of AC-DILI in each group.

AC is widely used for short-course treatment, and therefore the clinical application of this body of data is not clear at present. In the 2011 GWAS of AC-DILI cases, Lucena et al. estimated the positive predictive value to be 0.1% for carriage of both HLA-A*02:01 and DQB1*06:02 in Northwestern Europeans and 0.13% for carriage of both HLA-B*18:01 and DQB1*06:02 among Spanish populations. The strongest negative predictive value in this analysis was seen for the absence of HLA-A*02:01 in Northwestern Europeans. Thus, because of the overall rarity of AC-induced DILI, HLA typing based on known associations does not currently meet thresholds to warrant routine screening prior to AC therapy [121].

Of note, some of the HLA associations identified as possible contributors to the pathogenesis of AC-DILI are also associated with other liver diseases of varying etiologies. For example, the DRB1*15:01-DQB1*06:02 haplotype is strongly associated with primary sclerosing cholangitis and weakly associated with primary biliary cirrhosis. This allele might also confer a protective effect in type 1 autoimmune hepatitis. The HLA-DRB1*15 family has also been associated with nitrofurantoin- and halothane-induced hepatitis as well as with allergy to ragweed pollen [179]. More work is needed to determine the exact mechanisms of HLA-DRB1*15 alleles in these diseases as well as in AC-mediated DILI.

21.6 STRATIFIED APPROACHES TO PREVENT DRUG HYPERSENSITIVITY AND THEIR TRANSLATION INTO CLINICAL MEDICINE

In general, diagnostic tests to identify the specific drug cause of an ADR, including DHS, are imprecise. The ability to use stratified approaches to identify patients at risk for any ADR via pharmacogenetic screening prior to drug administration is an attractive concept given the lack of perfect diagnostic tests and the substantial morbidity and cost associated with DHS.

21.6.1 Criteria for Stratified Approaches

The criteria discussed in the following subsections applies to any pharmacogenomic stratified treatment approach; however, there are specific issues with regard to immunogenetics and HLA testing that will be discussed. From a cost and quality assurance standpoint, the increasing availability of accurate and accessible methods with rapid turnaround times (e.g., single allele PCR-based assays) has made these tests accessible to routine diagnostic laboratories. It is advantageous for multiple criteria to be satisfied in order for stratified approaches to be cost-effective and feasible for implementation into routine clinical practice.

21.6.1.1 The DHS Under Investigation Must be Prevalent and Severe so as to Warrant Screening

If the ultimate goal of investigation is the incorporation of specific screening and/or therapeutics into clinical practice, research should be targeted at DHS that result in an unacceptable rate of mortality, morbidity, hospitalization, or therapeutic changes. Minor DHS (e.g., transient mild rash) are not appropriate targets given that the costs associated with research and testing will exceed the cost of current strategies for medical management. In addition, even if a mild rash without systemic symptoms is associated with a specific HLA allele, we know from clinical experience that the vast majority of patients can continue treatment with symptomatic management using antihistamines and/or topical steroids. The disappearance of the rash despite continued treatment in these patients has been presumed to be associated with immunological tolerance.

Other important considerations are that the specific drug under investigation should be widely used and efficacious, and have an overall favorable side-effect profile. Generally, alternative therapies should be limited, less efficacious, associated with negative side effects, or require costly screening or clinical/laboratory monitoring.

21.6.1.2 Clear Diagnostic Criteria Must be Used to Define Putative HLA-Associated ADRs

Our experience with AHR highlights the importance of diagnostic precision. As noted previously, reliance on clinical criteria alone and, in particular, retrospective data to diagnose AHR results in a high rate of false-positive results due to the nonspecific nature of common presenting symptoms. In fact, results from blinded comparative trials to assess the efficacy and safety of antiretroviral therapies that include abacavir-containing regimens show that up to 7% of AHR clinical diagnoses occur in patients who were not receiving abacavir [180–182]. The introduction of epicutaneous patch testing to specifically identify immunologically-mediated cases of AHR substantially reduce diagnostic error and provide a foundation upon which to decipher the immunogenetic correlates of AHR [51,70].

It should be noted that the sensitivity of patch testing for any given drug varies according to clinical phenotype of the DHS and the implicated drug. Generally, the sensitivity of patch testing is less than 70%, although higher sensitivities have been noted for AHR (for which the diagnostic sensitivity as calculated from the PREDICT-1 study is estimated to be 87% [68]. Patch testing has also been associated with a higher sensitivity of carbamazepine-DIHS/DRESS/HSS and AGEP and a lower sensitivity for epidermal diseases such as FDE and SJS/TEN [51,183]. The lower sensitivity of patch testing for drug-induced FDE and SJS/TEN may reflect the presence of resident T cells or the characteristics of the compartment in which these T-cells (epidermal) versus the drug (dermal) reside, as well as the presence or absence of specific drug-metabolizing enzymes in the skin [184].

Additionally, a number of *in vitro* laboratory assays have been employed to complement clinical diagnosis and patch testing in an effort to define immune mechanisms of ADRs. These include lymphocyte transformation assays, in which T cells from hypersensitive patients are collected and induced to proliferate in the presence of the implicated drug, as well as *ex vivo* assays that measure T-cell function following drug exposure, such as enzyme-linked immunospot (ELISpot) and intracellular cytokine release assays (Rive et al. [34]). While these assays are of some collective use in helping to support the diagnosis of an immune-mediated ADR, their sensitivity remains less than 100% and they have no role in pre-therapy screening because they rely on T-cell sensitization to the drug in question prior to testing. Also, because of the severity of the drug reactions involved, a negative diagnostic test is never used as the sole basis for rechallenge to any drug. Given the 55% PPV of HLA-B*57:01 for abacavir hypersensitivity, 45% of patients positive for HLA-B*57:01 are expected to tolerate abacavir. It is of interest, therefore, that abacavir treatment of PBMCs isolated from healthy, abacavir-unexposed healthy blood donors results in a CD8+ T-cell response in 100% of 10–14 day cultures [45]. This is in contrast to abacavir-specific patch test studies, which are consistently negative in HLA-B*57:01 positive, abacavir-naïve patients [185].

21.6.1.3 The Screening Test Must Have Sufficient Predictive Value to Ensure That a Positive (or Negative) Result Accurately Reflects the Likelihood of an ADR

Although multiple HLA alleles have been reported in relation to DHS, only a handful of these reports demonstrate odds ratios that strongly favor association. Importantly, a high negative predictive value (NPV) is critical to ensure drug safety. Additionally, the number needed to test (NNT) to prevent one case of DHS must be appropriately low to allow for cost- and resource-effective screening across a population. For some drug reactions, the combination of a prevalent HLA allele and the low prevalence of the DHS can mean a much lower positive predictive value for the screening test and an impractical number of cases needed to test to prevent one case of DHS (e.g., flucloxacillin-associated DILI).

21.6.1.4 The Screening Test Must Maintain Adequate Predictive Values Across a Defined Population of Patients

A corollary to this criterion is that the HLA allele must be sufficiently prevalent in the population under study. For example, as discussed previously, the presence of HLA-B*15:02 is strongly associated with risk of carbamazepine-induced SJS/TEN in persons of Asian descent, owing to the relatively high prevalence of this allele in this population (10–15% in Han Chinese; NPV 100%, PPV 1–4%, NNT 461 [153]. These results are not, however, generalizable to Caucasian populations in which the rate of HLA-B*15:02 carriage is substantially lower, as shown in Figure 21.5 (allele prevalence <1% in Caucasians; estimated NPV <100%, NNT 10,000) [34,153,186]. Conversely, the HLA-B*57:01 association with abacavir hypersensitivity is sufficiently robust so as to warrant screening of all patients prior to initiation of therapy, regardless of ethnic background (NPV 100%, PPV 55%, NNT 13 across populations) [34,68,69,187].

The question remains as to whether or not HLA-B*57:01 screening is cost-effective in certain populations in Africa and part of Southeast Asia, where the prevalence of HLA-B*57:01 carriage is <1% and there exists little migration or population admixture. Thus, careful definition of the target screening population ensures that the test maintains adequate predictive value in the clinical setting. It should be noted that stratification of patients based on race or reported ethnic background must be approached with some degree of caution given that patient phenotype does not always represent genetics, especially in populations with a high degree of genetic admixture.

21.6.1.5 The Screening Test Must be Available in Clinical Laboratories, Reproducible, Efficient, and Cost-Effective

Currently, high-resolution sequence-based typing is the gold standard for determining an individual HLA genotype. This method, while highly accurate, requires substantial technical infrastructure and expertise; until recently, it was primarily restricted to the research setting and tissue-typing laboratories at organ transplant centers. The demonstration that prospective HLA-B*57:01 screening is effective for the prevention of AHR, along with the subsequent recommendation for universal screening, has increased the market for HLA-typing resources, and multiple commercial laboratories are now performing allele-specific PCR-based tests in the United States and Europe, and a number of rapid and cost-effective assays now exist that deliver results within an acceptable turnaround time for the purposes of initiating antiretroviral therapy. Several technologies have been proposed to improve the cost and efficiency of HLA-B*57:01 identification, many of which have already been applied to HLA typing in other contexts.

HLA-B*57:01 belongs to the B17 family of alleles that also includes HLA-B*58 and at least nine different HLA-B*57 subtypes (HLA-B*57:01 through B*57:09). Within the B17 family, the HLA-B*58:01, HLA-B*57:02, and HLA-B*57:03 alleles are prevalent among diverse populations and share significant homology with HLA-B*57:01, varying only at 1–4 amino acids in the peptide binding groove. Notably, HLA-B*57:01 is the only allotype associated with AHR, and thus HLA screening in this context must discriminate between HLA-B*57:01 and these closely related alleles. In addition to high-resolution genotyping, PCR-based diagnostics have been developed to address this problem. Sequence-specific primer PCR (SSP-PCR) employs primers designed to amplify HLA-B*57:01 sequences but not closely related B17 sequences with high specificity (100%), sensitivity (99.4%), and reproducibility across multiple laboratories [188,189]. These protocols have also been modified to include a melting curve analysis for rapid identification of SSP-PCR products with equivalent accuracy to that of sequence-based assays (100% specificity and sensitivity) [190].

Although these PCR-based protocols currently require DNA extraction from whole blood prior to analysis, it is possible to batch samples in order to decrease costs and turnaround time. Additionally, efforts to optimize PCR parameters for use on whole blood and saliva samples are underway and, if successful, might markedly simplify and standardize these protocols for use in clinical laboratories [191,192].

One strategy to improve the cost and efficiency of PCR-based diagnostics employs serologic screening for any HLA-B*57 or B*58 allele prior to molecular testing. In 2006, Martin et al. demonstrated that flow cytometry using commercially available anti-B17 antibodies and whole blood samples from the HIV clinic accurately identified carriers of B17 alleles with 100% sensitivity. While this method does not differentiate HLA-B*57:01 from other B17 alleles, identification of individuals who do not carry a B17 allele would afford the physician the opportunity to safely start abacavir therapy in a timely manner, and would reduce the number of samples for which HLA tying is indicated. Martin et al. estimated that serologic stratification decreases the number of samples referred for molecular testing to approximately 10% of the total and therefore substantially reduces cost [193]. Similarly, Kostenko et al. developed an anti-B17 monoclonal antibody, 3E12, and demonstrated that labeled 3E12 can be incorporated into multicolor FACS assays that are currently used for the routine assessment of T-cell subsets (CD4+ T-cell count) in HIV patients. This would provide information regarding B17 status without modification of laboratory workflow. Samples that test positive by this method would then undergo HLA genotyping to determine if the patient is positive for HLA-B*57:01 [194].

Alternative proposed strategies for HLA-B*57:01 screening have been investigated. These include identification of SNPs known to be in strong LD with HLA-B*57:01. The HCP5 SNP (rs2395029) was evaluated as a candidate for such analysis, but studies demonstrating incomplete concordance with HLA-B*57:01 across populations raise concerns about the utility of this marker in clinical practice as it would have a <100% negative predictive value [36,195]. Additionally, inspection of HIV-1 viral sequences for mutations associated with selection by HLA-B*57:01 cytotoxic T lymphocytes has also been evaluated as a potential screening tool. However, the strength of these associations was found not to be sufficient for introduction into clinical practice [196,197].

21.6.1.6 A Robust Clinical and Quality Control Infrastructure Must be in Place to Ensure Appropriate Test Administration and Interpretation of Results

Successful application of HLA screening tests that meet the criteria just given requires that the test be administered according to standardized protocols and that clinicians be able to accurately interpret test results to inform clinical decision making. Importantly, if widespread pharmacogenomic screening is to be introduced into clinical practice, consistency among commercially available products and testing facilities will need to be ensured.

21.6.2 Current Recommendations for HLA Screening

Despite numerous reports of HLA-associated drug toxicities, only work involving HLA-B*57:01 screening for AHR has been generally applied in a routine clinical setting based on evidence based medicine and treatment guidelines. It is currently recommended that HLA-B*15:02 screening for carbamazepine SJS/TEN be selectively applied to those populations at higher risk such as South-east Asians; moreover, HLA-B*58:01 screening for allopurinol SJS/TEN and/or DRESS/DIHS/HSS is still under study/consideration in most jurisdictions. The technology exists to create quality-assured low-cost tests with reasonable turnaround times. However, for most drugs there are still many hurdles to overcome prior to their routine testing in clinical practice. The notable success with abacavir—the high prevalence of the DHS, the 100% negative predictive value of HLA-B*57:01 for AHR, the high positive predictive value, the generalizability across ethnicities, and the high prevalence of allele carriage in Caucasian populations combined with convergent clinical and scientific evidence to support the exquisite specificity of HLA-B*57:01 for AHR—has created a roadmap from discovery to translation. Following the PREDICT-1 and SHAPE trials in 2008, both of which demonstrated the efficacy of HLA-B*57:01 screening for the prevention of AHR, the FDA issued a black-box warning against the use of abacavir in patients known to carry the HLA-B*57:01 allele. Multiple groups, including the US FDA and Department of Health and Human Services, the European Medicines Agency, the Clinical Pharmacogenetics Implementation Consortium, and international HIV/AIDS organizations, now recommend universal screening for HLA-B*57:01 carriage in any patient for which abacavir therapy is under consideration, and exclusion of abacavir therapy for any patient with a positive test [187,198-201]. In resource-limited settings where HLA-B*57:01 genotyping is not available and its prevalence in the population is near zero, some groups have advocated for initiation of abacavir therapy with close clinical monitoring for symptoms of AHR and its prompt discontinuation if these symptoms develop [187].

Similarly, the discovery of the HLA-B*15:02 association with carbamazepine-induced SJS/TEN in Asian populations prompted the FDA in 2007 to recommend genetic screening in persons of Asian ancestry prior to initiation of carbamazepine [202]. This recommendation has since been modified to include persons “at risk” of carbamazepine-induced SJS/TEN as determined by the prevalence of HLA-B*15:02 carriage in the population (Figure 21.5). Specifically, screening is recommended for persons of Han Chinese descent and those originating from Vietnam, Cambodia, the Reunion Islands, Thailand, India, Malaysia, and Hong Kong. Further, if a patient is found to carry the HLA-B*15:02 gene, regardless of ancestry, carbamazepine therapy is contraindicated [186]. Additionally, multiple studies have demonstrated that not all carbamazepine-induced SJS/TEN can be attributed to HLA-B*15:02 carriage (Table 21.2) and thus careful clinical monitoring is recommended for all patients treated with carbamazepine regardless of HLA genotype [153,186]. The strong association of allopurinol induced

DRESS/DIHS/HS/SJS/TEN with HLA-B*58:01 is now under consideration by many jurisdictions including the Taiwanese Bureau of National Health Insurance for recommendation for routine HLA-B*58:01 genotyping prior to allopurional prescription. The FDA has currently not made a recommendation for HLA-B*58:01 screening prior to allopurinol prescription. A large Taiwanese clinical trial investigating whether prospective HLA-B*58:01 screening will reduce the incidence of allopurinol-associated ADRs is currently underway [203]. For other drugs (e.g., flucloxacillin) screening appears neither cost-effective nor feasible—almost 14,000 patients would need to be screened to prevent one case of flucloxacillin-induced DILI. Other stratified screening approaches exist, such as HLA-B*57:01 testing in patients who have already developed suspected flucloxacillin-DILI cases in populations with a high prevalence of HLA-B*57:01 carriage (e.g., Caucasians). However, there is no current base of evidence to support this strategy.

21.7 THE FUTURE DIRECTION OF STRATIFIED MEDICINE IN DRUG-INDUCED HYPERSENSITIVITY

The translation of HLA-B*57:01 from its discovery to a routine test for it in HIV clinics to identify persons at risk of AHR and exclude them from abacavir treatment has created a roadmap for successful implementation of pharmacogenomic stratified medicine. Several challenges exist, however, when trying to replicate and fully execute this road map for other HLA-associated drug reactions. In particular the low positive predictive values, the fewer than 100% negative predictive values, and the lack of generalizability across different populations for many HLA-drug associations create concerns not only regarding lack of feasibility and cost-effectiveness, but more broadly in relation to patient safety if screening lead to a false sense of security and lowers clinical vigilance.

It is important that due diligence be applied to the study and interpretation of pharmacogenomic research on drug hypersensitivity. Newer approaches (e.g., GWAS and whole-exome sequencing) are now widely available, but they currently do not provide read lengths long enough to derive high-resolution sequence-based HLA typing. In addition, given LD patterns, caution should be applied when trying to impute an HLA association from a SNP or SNP pattern; moreover, SNP associations in isolation do not give sufficient information about the immunogenetic basis of DHS. For this reason, ancestral haplotype mapping of the region at risk remains the gold standard.

Continuing research promises to further define factors leading to the incomplete positive predictive value of a specific HLA class I or II allele or an HLA class I/II pairing for a drug syndrome. These likely comprise factors outside of the MHC and may lead to further clinical utility for stratified pharmacogenomic screening approaches. More aptly, however, they will provide great insights into the immunopathogenesis of drug hypersensitivity and other inflammatory diseases. The abacavir example has significantly furthered our understanding of the science of drug hypersensitivity and provided proof of concept for the altered peptide model. More recent evidence suggests that this model may extend to phenotypically distinct immunologically mediated adverse drug reactions (e.g., carbamazepine-associated SJS/TEN). Thus, it may be more broadly applicable to immunologically mediated ADRs and in particular T-cell–mediated syndromes where non-covalent interactions between the MHC and the drug are likely. Quite remarkably, it is possible that, through the utilization of currently available structural, biochemical, and functional approaches, the interaction between HLA-B*57:01 and abacavir, as well as the potential for the development of an HLA-B*57:01-restricted CD8+ T-cell–dependent reaction, may have been defined prior to this drug being used in humans. For reasons such as this, exciting future applications of stratified pharmacogenomics will include approaches to pre-clinically define interactions between HLA and drugs using a combination of structural, biochemical, and functional approaches like those recently defined for abacavir. Such approaches might be applied pre-clinically, at the time of initial drug development and design, before the use of drugs in humans as well as at the first sign of any ADR in the pre-marketing or early post-marketing phase of drug development.

Key Concepts in HLA and the Pharmacogenomics of Drug Hypersensitivity

- The HLA genes, in particular HLA-B, are the most polymorphic genes in the human genome. They encode the HLA class I and HLA class II proteins that bind and present peptide antigens to T cells. The interaction of T cells with the HLA-peptide complex elicits a T-cell immune response; thus, the HLA proteins are central to normal immune function. The specificity of HLA-peptide binding is determined by the amino acid composition of the HLA peptide binding groove. A growing body of data shows that HLA-restricted T-cell–mediated immune responses are involved in many adverse drug reactions.
- Multiple models have been proposed to describe the mechanisms of HLA-associated adverse drug reactions. These include:
  - The hapten/prohapten model, which is based on the concept that a drug or drug metabolite may covalently bind to an endogenous peptide and thereby create a neo-antigen that is recognized by T cells to trigger an immune reaction.
  - The p-i model, which hypothesizes that a non-covalent interaction between a drug and the HLA protein can directly activate T cells in a peptide-independent manner.
  - The altered peptide repertoire model, which posits that a drug may non-covalently bind to the peptide binding groove

of an HLA molecule to change the chemistry and shape of the peptide binding cleft. This results in the selection of a different repertoire of endogenous peptides from that selected in the absence of the drug. This altered repertoire includes novel antigens that are recognized by T cells to elicit an immune response. Strong evidence suggests that the altered peptide repertoire model is central to the pathogenesis of phenotypically distinct syndromes (e.g., abacavir hypersensitivity and carbamazepine-associated SJS/TEN).

- The use of high-resolution HLA genotyping is the current gold standard method for determining HLA associations in drug hypersensitivity reactions. Current research suggests that it is the HLA allele that is necessary but not sufficient for the development of hypersensitivity. Alternative methods include identification SNPs that are LD with HLA genes and their use as surrogate markers for HLA carriage. HLA associations defined by this approach, however, must be interpreted with caution in instances where the LD between the tag SNP and the HLA gene of interest is incomplete.
  High-resolution HLA typing with ancestral haplotype mapping is the current gold standard in specifically defining drug-HLA associations.
- Many associations have now been described between various T-cell–mediated DHS and class I and/or class II HLA alleles. In the case of the association of AHR and HLA-B*57:01, this has now been translated from discovery to a routine screening test in clinical practice and has created a roadmap for other drug–HLA associations. Even though most drugs are associated with several innate hurdles that are nonmodifiable, making translation infeasible, they provide many additional valuable insights into the immunopathogenesis of these reactions. For some drugs, such as carbamazepine, regulatory agencies such as the FDA have mandated changes to product labels to recommend risk stratification and screening of select populations at high risk for HLA-B*15:02 carriage prior to carbamazepine prescription. In the future, heightened understanding of the specific structural and functional interactions with drugs and the MHC will be key to developing strategies to preclinically predict drugs at risk for causing severe immunologically mediated ADRs and to inform safer drug development and design.

ACRONYMS

ADR	adverse drug reaction
AGEP	acute generalized exanthematous pustulosis
AHR	abacavir hypersensitivity reaction
DHS	drug hypersensitivity syndrome
DIHS	drug-induced hypersensitivity syndrome
DILI	drug-induced liver injury
DRESS	drug reaction with eosinophilia and systemic symptoms
GWAS	genome-wide association study
HLA	human leukocyte antigen
HSS	hypersensitivity syndrome
LD	linkage disequilibrium
MHCI	major histocompatibility complex class I
MHCII	major histocompatibility complex class II
MPE	maculopapular eruption
NNT	number needed to test/treat
NPV	negative predictive value
NVP	nevirapine
OR	odds ratio
PCR	polymerase chain reaction
PPV	positive predictive value
SCAR	severe cutaneous adverse reaction
SJS	Stevens-Johnson syndrome
SNP	single-nucleotide polymorphism
TCR	T-cell receptor
TEN	toxic epidermal necrolysis

REFERENCES

[1] Roujeau JC, Stern RS. Severe adverse cutaneous reactions to drugs. N Engl J Med 1994;331(19):1272–85.

[2] Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. Ann Pharmacother 2008;42(7):1017–25.

[3] Hakkarainen KM, Hedna K, Petzold M, Hagg S. Percentage of patients with preventable adverse drug reactions and preventability of adverse drug reactions–a meta-analysis. PLoS ONE 2012;7(3):e33236.

[4] Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;279(15):1200–5.

[5] Suh DC, Woodall BS, Shin SK, Hermes-De Santis ER. Clinical and economic impact of adverse drug reactions in hospitalized patients. Ann Pharmacother 2000;34(12):1373–9.

[6] Pirmohamed M, Friedmann PS, Molokhia M, Loke YK, Smith C, Phillips E, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 2004;329(7456):15–19.

[7] Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. JAMA 1997;277(4):301–6.

[8] Bates DW, Spell N, Cullen DJ, Burdick E, Laird N, Petersen LA, et al. The costs of adverse drug events in hospitalized patients. Adverse drug events prevention study group. JAMA 1997;277(4):307–11.

[9] Bharadwaj M, Illing P, Theodossis A, Purcell AW, Rossjohn J, McCluskey J. Drug hypersensitivity and human leukocyte antigens of the major histocompatibility complex. Annu Rev Pharmacol Toxicol 2012;52:401–31.

[10] Pavlos R, Mallal S, Phillips E. HLA and pharmacogenetics of drug hypersensitivity. Pharmacogenomics 2012;13(11):1285–306.

[11] Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. Annu Rev Immunol 2006;24:419–66.

[12] Horton R, Wilming L, RV, Lovering RC, Bruford EA, Khodiyar VK, et al. Gene map of the extended human MHC. Nat Rev Genet 2004;5(12):889–99.

[13] Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. BMC Immunol 2008;9:1.

[14] Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. Structure of the human class I histocompatibility antigen, HLA-A2. Nature 1987;329(6139):506–12.

[15] Macdonald WA, Purcell AW, Mifsud NA, Ely LK, Williams DS, Chang L, et al. A naturally selected dimorphism within the HLA-B44 supertype alters class I structure, peptide repertoire, and T-cell recognition. J Exp Med 2003;198(5):679–91.

[16] van Kaer L. Major histocompatibility complex class I-restricted antigen processing and presentation. Tissue Antigens 2002;60(1):1–9.

[17] Wearsch PA, Cresswell P. Selective loading of high-affinity peptides onto major histocompatibility complex class I molecules by the tapasin-ERp57 heterodimer. Nat Immunol 2007;8(8):873–81.

[18] Chen W, Khilko S, Fecondo J, Margulies DH, McCluskey J. Determinant selection of major histocompatibility complex class I-restricted antigenic peptides is explained by class I-peptide affinity and is strongly influenced by nondominant anchor residues. J Exp Med 1994;180(4):1471–83.

[19] Starr TK, Jameson SC, Hogquist KA. Positive and negative selection of T cells. Annu Rev Immunol 2003;21:139–76.

[20] Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 1993;364(6432):33–9.

[21] Painter CA, Stern LJ. Conformational variation in structures of classical and non-classical MHCII proteins and functional implications. Immunol Rev 2012;250(1):144–57.

[22] Madden DR, Gorga JC, Strominger JL, Wiley DC. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. Nature 1991;353(6342):321–5.

[23] Prugnolle F, Manica A, Charpentier M, Guegan JF, Guernier V, Balloux F. Pathogen-driven selection and worldwide HLA class I diversity. Curr Biol 2005;15(11):1022–7.

[24] Qutob N, Balloux F, Raj T, Liu H, Marion de Procé S, Trowsdale J, et al. Signatures of historical demography and pathogen richness on MHC class I genes. Immunogenetics 2012;64(3):165–75.

[25] Hraber P, Kuiken C, Yusim K. Evidence for human leukocyte antigen heterozygote advantage against hepatitis C virus infection. Hepatology 2007;46(6):1713–21.

[26] Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, et al. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. Science 1999;283(5408):1748–52.

[27] Thursz MR, Thomas HC, Greenwood BM, Hill AV. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. Nat Genet 1997;17(1):11–12.

[28] Dawkins R, Leelayuwat C, Gaudieri S, Tay G, Hui J, Cattley S, et al. Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease. Immunol Rev 1999;167:275–304.

[29] Archbold JK, Macdonald WA, Gras S, Ely LK, Miles JJ, Bell MJ, et al. Natural micropolymorphism in human leukocyte antigens provides a basis for genetic control of antigen recognition. J Exp Med 2009;206(1):209–19.

[30] Burrows JM, Wynn KK, Tynan FE, Archbold J, Miles JJ, Bell MJ, et al. The impact of HLA-B micropolymorphism outside primary peptide anchor pockets on the CTL response to CMV. Eur J Immunol 2007;37(4):946–53.

[31] Gao X, Nelson GW, Karacki P, Martin MP, Phair J, Kaslow R, et al. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. N Engl J Med 2001;344(22):1668–75.

[32] D'Orsogna LJ, Roelen DL, Doxiadis II, Claas FH. TCR cross-reactivity and allorecognition: new insights into the immunogenetics of allorecognition. Immunogenetics 2012;64(2):77–85.

[33] Kapur V, Maffei JT, Greer RS, Li LL, Adams GJ, Musser JM. Vaccination with streptococcal extracellular cysteine protease (interleukin-1 beta convertase) protects mice against challenge with heterologous group A streptococci. Microb Pathog 1994;16(6):443–50.

[34] Rive CM, Bourke J, Phillips EJ. Testing for drug hypersensitivity syndromes. Clin Biochem Rev—Aust Assoc Clin Biochem 2013;34(1):15–38.

[35] Gaudieri S, Dawkins RL, Habara K, Kulski JK, Gojobori T. SNP profile within the human major histocompatibility complex reveals an extreme and interrupted level of nucleotide diversity. Genome Res 2000;10(10):1579–86.

[36] Melis R, Lewis T, Millson A, Lyon E, McMillin GA, Slev PR, et al. Copy number variation and incomplete linkage disequilibrium interfere with the HCP5 genotyping assay for abacavir hypersensitivity. Genet Test Mol Biomarkers 2012;16(9):1111–4.

[37] Eshki M, Allanore L, Musette P, Milpied B, Grange A, Guillaume JC, et al. Twelve-year analysis of severe cases of drug reaction with eosinophilia and systemic symptoms: a cause of unpredictable multiorgan failure. Arch Dermatol 2009;145(1):67–72.

[38] Shiohara T, Inaoka M, Kano Y. Drug-induced hypersensitivity syndrome (DIHS): a reaction induced by a complex interplay among herpesviruses and antiviral and antidrug immune responses. Allergol Int: J Jpn Soc Allergol 2006;55(1):1–8.

[39] Tohyama M, Hashimoto K, Yasukawa M, Kimura H, Horikawa T, Nakajimas K, et al. Association of human herpesvirus 6 reactivation with the flaring and severity of drug-induced hypersensitivity syndrome. Br J Dermatol 2007;157(5):934–40.

[40] Pirmohamed M, Friedmann PS, Molokhia M, Loke YK, Smith C, Phillips E, et al. Phenotype standardization for immune-mediated drug-induced skin injury. Clin Pharmacol Ther 2011;89(6):896–901.

[41] Andrews E, Daly AK. Flucloxacillin-induced liver injury. Toxicology 2008;254(3):158–63.

[42] Daly AK, Day CP. Genetic association studies in drug-induced liver injury. Drug Metab Rev 2012;44(1):116–26.

[43] Daly AK, Day CP. Genetic association studies in drug-induced liver injury. Semin Liver Dis 2009;29(4):400–11.

[44] Schnyder B, Mauri-Hellweg D, Zanni M, Bettens F, Pichler WJ. Direct, MHC-dependent presentation of the drug sulfamethoxazole to human alphabeta T-cell clones. Eur J Clin Invest 1997;100(1):136–41.

[45] Chessman D, Kostenko L, Lethborg T, Purcell AW, Williamson NA, Chen Z, et al. Human leukocyte antigen class I-restricted activation of CD8+ T cells provides the immunogenetic basis of a systemic drug hypersensitivity. Immunity 2008;28(6):822–32.

[46] Naisbitt DJ, Britschgi M, Wong G, Farrell J, Depta JP, Chadwick DW, et al. Hypersensitivity reactions to carbamazepine: characterization of the specificity, phenotype, and cytokine profile of drug-specific T-cell clones. Mol Pharmacol 2003;63(3):732–41.

[47] Wei CY, Chung WH, Huang HW, Chen YT, Hung SI. Direct interaction between HLA-B and carbamazepine activates T cells in patients with Stevens-Johnson syndrome. J Allergy Clin Immunol 2012;129(6):1562–9. e5

[48] Pichler W, Yawalkar N, Schmid S, Helbling A. Pathogenesis of drug-induced exanthems. Allergy 2002;57(10):884–93.

[49] Hari Y, Frutig-Schnyder K, Hurni M, Yawalkar N, Zanni MP, Schnyder B, et al. T-cell involvement in cutaneous drug eruptions. Clin Exp Allergy 2001;31(9):1398–408.

[50] Le Cleach L, Delaire S, Boumsell L, Bagot M, Bourgault-Villada I, Bensussan A, et al. Blister fluid T lymphocytes during toxic epidermal necrolysis are functional cytotoxic cells which express human natural killer (NK) inhibitory receptors. Clin Exp Immunol 2000;119(1):225–30.

[51] Phillips EJ, Sullivan JR, Knowles SR, Shear NH. Utility of patch testing in patients with hypersensitivity syndromes associated with abacavir. AIDS 2002;16(16):2223–5.

[52] Pichler WJ. Delayed drug hypersensitivity reactions. Ann Intern Med 2003;139(8):683–93.

[53] Padovan E, Mauri-Hellweg D, Pichler WJ, Weltzien HU. T-cell recognition of penicillin G: structural features determining antigenic specificity. Eur J Immunol 1996;26(1):42–8.

[54] Naisbitt DJ, Gordon SF, Pirmohamed M, Burkhart C, Cribb AE, Pichler WJ, et al. Antigenicity and immunogenicity of sulphamethoxazole: demonstration of metabolism-dependent haptenation and T-cell proliferation *in vivo*. Br J Pharmacol 2001;133(2):295–305.

[55] Ostrov DA, Grant BJ, Pompeu YA, Sidney J, Harndahl M, Southwood S, et al. Drug hypersensitivity caused by alteration of the MHC-presented self-peptide repertoire. Proc Nat Acad Sci USA 2012;109(25):9959–64.

[56] Zanni MP, von Greyerz S, Schnyder B, Brander KA, Frutig K, Hari Y, et al. HLA-restricted, processing- and metabolism-independent pathway of drug recognition by human alpha beta T lymphocytes. Eur J Clin Invest 1998;102(8):1591–8.

[57] Zanni MP, von Greyerz S, Schnyder B, Wendland T, Pichler WJ. Allele-unrestricted presentation of lidocaine by HLA-DR molecules to specific alphabeta+ T-cell clones. Int Immunol 1998;10(4):507–15.

[58] Burkhart C, von Greyerz S, Depta JP, Naisbitt DJ, Britschgi M, Park KB, et al. Influence of reduced glutathione on the proliferative response of sulfamethoxazole-specific and sulfamethoxazole-metabolite-specific human CD4+ T-cells. Br J Pharmacol 2001;132(3):623–30.

[59] Pichler WJ, Beeler A, Keller M, Lerch M, Posadas S, Schmid D, et al. Pharmacological interaction of drugs with immune receptors: the p-i concept. Allergol Int: J Jpn Soc Allergol 2006;55(1):17–25.

[60] Norcross MA, Luo S, Lu L, Boyne MT, Gomarteli M, Rennels AD, et al. Abacavir induces loading of novel self-peptides into HLA-B*57: 01: an autoimmune model for HLA-associated drug hypersensitivity. AIDS 2012;26(11):F21–9.

[61] Illing PT, Vivian JP, Dudek NL, Kostenko L, Chen Z, Bharadwaj M, et al. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. Nature 2012;486(7404):554–8.

[62] Cutrell AG, Hernandez JE, Fleming JW, Edwards MT, Moore MA, Brothers CH, et al. Updated clinical risk factor analysis of suspected hypersensitivity reactions to abacavir. Ann Pharmacother 2004;38(12):2171–2.

[63] Shapiro M, Ward KM, Stern JJ. A near-fatal hypersensitivity reaction to abacavir: case report and literature review. AIDS Read 2001;11(4):222–6.

[64] Symonds W, Cutrell A, Edwards M, Steel H, Spreen B, Powell G, et al. Risk factor analysis of hypersensitivity reactions to abacavir. Clin Ther 2002;24(4):565–73.

[65] Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, et al. Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. Lancet 2002;359(9308):727–32.

[66] Hetherington S, Hughes AR, Mosteller M, Shortino D, Baker KL, Spreen W, et al. Genetic variations in HLA-B region and hypersensitivity reactions to abacavir. Lancet 2002;359(9312):1121–2.

[67] Cao K, Hollenbach J, Shi X, Shi W, Chopek M, Fernandez-Vina MA. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. Hum Immunol 2001;62(9):1009–30.

[68] Mallal S, Phillips E, Carosi G, Molina JM, Workman C, Tomazic J, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008;358(6):568–79.

[69] Saag M, Balu R, Phillips E, Brachman P, Martorell C, Burman W, et al. High sensitivity of human leukocyte antigen-b*5701 as a marker for immunologically confirmed abacavir hypersensitivity in white and black patients. Clin Infect Dis 2008;46(7):1111–8.

[70] Phillips EJ, Wong GA, Kaul R, Shahabi K, Nolan DA, Knowles SR, et al. Clinical and immunogenetic correlates of abacavir hypersensitivity. AIDS 2005;19(9):979–81.

[71] Adam J, Eriksson KK, Schnyder B, Fontana S, Pichler WJ, Yerly D. Avidity determines T-cell reactivity in abacavir hypersensitivity. Eur J Immunol 2012;42(7):1706–16.

[72] Chan SH, Tan T. HLA and allopurinol drug eruption. Dermatologica 1989;179(1):32–3.

[73] Hung SI, Chung WH, Liou LB, Chu CC, Lin M, Huang HP, et al. HLA-B*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. Proc Nat Acad Sci USA 2005;102(11):4134–9.

[74] Tassaneeyakul W, Jantararoungtong T, Chen P, Lin PY, Tiamkao S, Khunarkornsiri U, et al. Strong association between HLA-B*5801 and allopurinol-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in a Thai population. Pharmacogenet Genomics 2009;19(9):704–9.

[75] Kang HR, Jee YK, Kim YS, Lee CH, Jung JW, Kim SH, et al. Positive and negative associations of HLA class I alleles with allopurinol-induced SCARs in Koreans. Pharmacogenet Genomics 2011;21(5):303–7.

[76] Genin E, Schumacher M, Roujeau JC, Naldi L, Liss Y, Kazma R, et al. Genome-wide association study of Stevens-Johnson syndrome and toxic epidermal necrolysis in Europe. Orphanet J Rare Dis 2011;6:52.

[77] Somkrua R, Eickman EE, Saokaew S, Lohitnavy M, Chaiyakunapruk N. Association of HLA-B*5801 allele and allopurinol-induced Stevens Johnson syndrome and toxic epidermal necrolysis: a systematic review and meta-analysis. BMC Med Genet 2011;12:118.

[78] Chung WH, Hung SI, Hong HS, Hsih MS, Yang LC, Ho HC, et al. Medical genetics: a marker for Stevens-Johnson syndrome. Nature 2004;428(6982):486.

[79] Kulkantrakorn K, Tassaneeyakul W, Tiamkao S, Jantararoungtong T, Prabmechai N, Vannaprasaht S, et al. HLA-B*1502 strongly predicts carbamazepine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in Thai patients with neuropathic pain. Pain Pract: J World Inst Pain 2012;12(3):202–8.

[80] Then SM, Rani ZZ, Raymond AA, Ratnaningrum S, Jamal R. Frequency of the HLA-B*1502 allele contributing to carbamazepine-induced hypersensitivity reactions in a cohort of Malaysian epilepsy patients. Asian Pac J Allergy Immunol 2011;29(3):290–3.

[81] Man CB, Kwan P, Baum L, Yu E, Lau KM, Cheng AS, et al. Association between HLA-B*1502 allele and antiepileptic drug-induced cutaneous reactions in Han Chinese. Epilepsia 2007;48(5):1015–8.

[82] Wang Q, Zhou JQ, Zhou LM, Chen ZY, Fang ZY, Chen SD, et al. Association between HLA-B*1502 allele and carbamazepine-induced severe cutaneous adverse reactions in Han people of Southern China mainland. Seizure 2011;20(6):446–8.

[83] Wu XT, Hu FY, An DM, Yan B, Jiang X, Kwan P, et al. Association between carbamazepine-induced cutaneous adverse drug reactions and the HLA-B*1502 allele among patients in Central China. Epilepsy Behav 2010;19(3):405–8.

[84] Zhang Y, Wang J, Zhao LM, Peng W, Shen GQ, Xue L, et al. Strong association between HLA-B*1502 and carbamazepine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in mainland Han Chinese patients. Eur J Clin Pharmacol 2011;67(9):885–7.

[85] Hung SI, Chung WH, Jee SH, Chen WC, Chang YT, Lee WR, et al. Genetic susceptibility to carbamazepine-induced cutaneous adverse drug reactions. Pharmacogenet Genomics 2006;16(4):297–306.

[86] Tassaneeyakul W, Tiamkao S, Jantararoungtong T, Chen P, Lin SY, Chen WH, et al. Association between HLA-B*1502 and carbamazepine-induced severe cutaneous adverse drug reactions in a Thai population. Epilepsia 2010;51(5):926–30.

[87] Locharernkul C, Loplumlert J, Limotai C, Korkij W, Desudchit T, Tongkobpetch S, et al. Carbamazepine and phenytoin induced Stevens-Johnson syndrome is associated with HLA-B*1502 allele in Thai population. Epilepsia 2008;49(12):2087–91.

[88] Chang CC, Too CL, Murad S, Hussein SH. Association of HLA-B*1502 allele with carbamazepine-induced toxic epidermal necrolysis and Stevens-Johnson syndrome in the multi-ethnic Malaysian population. Int J Dermatol 2011;50(2):221–4.

[89] Kim SH, Lee KW, Song WJ, Kim SH, Jee YK, Lee SM, et al. Carbamazepine-induced severe cutaneous adverse reactions and HLA genotypes in Koreans. Epilepsy Res 2011;97(1–2):190–7.

[90] Kaniwa N, Saito Y, Aihara M, Matsunaga K, Tohkin M, Kurose K, et al. HLA-B*1511 is a risk factor for carbamazepine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in Japanese patients. Epilepsia 2010;51(12):2461–5.

[91] Ikeda H, Takahashi Y, Yamazaki E, Fujiwara T, Kaniwa N, Saito Y, et al. HLA class I markers in Japanese patients with carbamazepine-induced cutaneous adverse reactions. Epilepsia 2010;51(2):297–300.

[92] McCormack M, Alfirevic A, Bourgeois S, Farrell JJ, Kasperaviciute D, Carrington M, et al. HLA-A*3101 and carbamazepine-induced hypersensitivity reactions in Europeans. N Engl J Med 2011;364(12):1134–43.

[93] Niihara H, Kakamu T, Fujita Y, Kaneko S, Morita E. HLA-A31 strongly associates with carbamazepine-induced adverse drug reactions but not with carbamazepine-induced lymphocyte proliferation in a Japanese population. J Dermatol 2012;39(7):594–601.

[94] Ozeki T, Mushiroda T, Yowang A, Takahashi A, Kubo M, Shirakata Y, et al. Genome-wide association study identifies HLA-A*3101 allele as a genetic risk factor for carbamazepine-induced cutaneous adverse drug reactions in Japanese population. Hum Mol Genet 2011;20(5):1034–41.

[95] Hung SI, Chung WH, Liu ZS, Chen CH, Hsih MS, Hui RC, et al. Common risk allele in aromatic antiepileptic-drug induced Stevens-Johnson syndrome and toxic epidermal necrolysis in Han Chinese. Pharmacogenomics 2010;11(3):349–56.

[96] Lin LC, Lai PC, Yang SF, Yang RC. Oxcarbazepine-induced Stevens-Johnson syndrome: a case report. Kaohsiung J Med Sci 2009;25(2):82–6.

[97] An DM, Wu XT, Hu FY, Yan B, Stefan H, Zhou D. Association study of lamotrigine-induced cutaneous adverse reactions and HLA-B*1502 in a Han Chinese population. Epilepsy Res 2010;92(2–3):226–30.

[98] Shi YW, Min FL, Liu XR, Zan LX, Gao MM, Yu MJ, et al. Hla-B alleles and lamotrigine-induced cutaneous adverse drug reactions in the Han Chinese population. Basic Clin Pharmacol Toxicol 2011;109(1):42–6.

[99] Kaniwa N, Sugiyama E, Saito Y, Kurose K, Maekawa, Hasegawa R, et al. Specific HLA types are associated with antiepileptic drug-induced Stevens-Johnson Syndrome and toxic epidermal necrolysis in Japanese subjects. Pharmacogenomics 2013;14(15):1821–31.

[100] Carr DF, Chaponda M, Jorgensen AL, Castro EC, van Oosterhout JJ, Khoo SH, et al. Association of human leukocyte antigen alleles and nevirapine hypersensitivity in a Malawian HIV-infected population. Clin Infect Dis 2013;56(9):1330–9.

[101] Lonjou C, Borot N, Sekula P, Ledger N, Thomas L, Halevy S, et al. A European study of HLA-B in Stevens-Johnson syndrome and toxic epidermal necrolysis related to five high-risk drugs. Pharmacogenet Genomics 2008;18(2):99–107.

[102] Kim SH, Kim M, Lee KW, Kim SH, Kang HR, Park HW, et al. HLA-B*5901 is strongly associated with methazolamide-induced Stevens-Johnson syndrome/toxic epidermal necrolysis. Pharmacogenomics 2010;11(6):879–84.

[103] Roujeau JC, Huynh TN, Bracq C, Guillaume JC, Revuz J, Touraine R. Genetic susceptibility to toxic epidermal necrolysis. Arch Dermatol 1987;123(9):1171–3.

[104] Kaniwa N, Saito Y, Aihara M, Matsunaga K, Tohkin M, Kurose K, et al. HLA-B locus in Japanese patients with anti-epileptics and allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis. Pharmacogenomics 2008;9(11):1617–22.

[105] Tohkin M, Kaniwa N, Saito Y, Sugiyama E, Kurose K, Nishikawa J, et al. A whole-genome association study of major determinants for allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis in Japanese patients. Pharmacogenomics J 2013;13(1):60–9.

[106] Chiu ML, Hu M, Ng MH, Yeung CK, Chan JC, Chang MM, et al. Association between HLA-B*58:01 allele and severe cutaneous adverse reactions with allopurinol in Han Chinese in Hong Kong. Br J Dermatol 2012;167(1):44–9.

[107] Martin AM, Nolan D, James I, Cameron P, Keller J, Moore C, et al. Predisposition to nevirapine hypersensitivity associated with HLA-DRB1*0101 and abrogated by low CD4 T-cell counts. AIDS 2005;19(1):97–9.

[108] Yuan J, Guo S, Hall D, Cammett AM, Jayadev S, Distel M, et al. Toxicogenomics of nevirapine-associated cutaneous and hepatic adverse events among populations of African, Asian, and European descent. AIDS 2011;25(10):1271–80.

[109] Littera R, Carcassi C, Masala A, Piano P, Serra P, Ortu F, et al. HLA-dependent hypersensitivity to nevirapine in Sardinian HIV patients. AIDS 2006;20(12):1621–6.

[110] Gatanaga H, Yazaki H, Tanuma J, Honda M, Genka I, Teruya K, et al. HLA-Cw8 primarily associated with hypersensitivity to nevirapine. AIDS 2007;21(2):264–5.

[111] Gao S, Gui XE, Liang K, Liu Z, Hu J, Dong B. HLA-dependent hypersensitivity reaction to nevirapine in Chinese Han HIV-infected patients. AIDS Res Hum Retroviruses 2012;28(6):540–3.

[112] Chantarangsu S, Mushiroda T, Mahasirimongkol S, Kiertiburanakul S, Sungkanuparph S, Manosuthi W, et al. HLA-B*3505 allele is

a strong predictor for nevirapine-induced skin adverse drug reactions in HIV-infected Thai patients. Pharmacogenet Genomics 2009;19(2):139–46.

[113] Alfirevic A, Jorgensen AL, Williamson PR, Chadwick DW, Park BK, Pirmohamed M. HLA-B locus in Caucasian patients with carbamazepine hypersensitivity. Pharmacogenomics 2006;7(6):813–8.

[114] Zhang FR, Liu H, Irwanto A, Fu XA, Li Y, Yu GQ, et al. HLA-B*13:01 and the dapsone hypersensitivity syndrome. N Eng J Med 2013;369(17):1620–8.

[115] Vitezica ZG, Milpied B, Lonjou C, Borot N, Ledger TN, Lefebvre A, et al. HLA-DRB1*01 associated with cutaneous hypersensitivity induced by nevirapine and efavirenz. AIDS 2008;22(4):540–1.

[116] Likanonsakul S, Rattanatham T, Feangvad S, Uttayamakul S, Prasithsirikul W, Tunthanathip P, et al. HLA-Cw*04 allele associated with nevirapine-induced rash in HIV-infected Thai patients. AIDS Res Ther 2009;6:22.

[117] Chantarangsu S, Mushiroda T, Mahasirimongkol S, Kiertiburanakul S, Sungkanuparph S, Manosuthi W, et al. Genome-wide association study identifies variations in 6p21.3 associated with nevirapine-induced rash. Clin Infect Dis 2011;53(4):341–8.

[118] Romano A, De Santis A, Romito A, Di Fonso M, Venuti A, Gasbarrini GB, et al. Delayed hypersensitivity to aminopenicillins is related to major histocompatibility complex genes. Ann Allergy Asthma Immunol 1998;80(5):433–7.

[119] Hautekeete ML, Horsmans Y, van Waeyenberge C, Demanet C, Henrion J, Verbist L, et al. HLA association of amoxicillin-clavulanate–induced hepatitis. Gastroenterology 1999;117(5):1181–6.

[120] O'Donohue J, Oien KA, Donaldson P, Underhill J, Clare M, MacSween RN, et al. Co-amoxiclav jaundice: clinical and histological features and HLA class II association. Gut 2000;47(5):717–20.

[121] Lucena MI, Molokhia M, Shen Y, Urban TJ, Aithal GP, Andrade RJ, et al. Susceptibility to amoxicillin-clavulanate-induced liver injury is influenced by multiple HLA class I and II alleles. Gastroenterology 2011;141(1):338–47.

[122] Singer JB, Lewitzky S, Leroy E, Yang F, Zhao X, Klickstein L, et al. A genome-wide study identifies HLA alleles associated with lumiracoxib-related liver injury. Nat Genet 2010;42(8):711–4.

[123] Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenomics J 2008;8(3):186–95.

[124] Daly AK, Aithal GP, Leathart JB, Swainsbury RA, Dang TS, Day CP. Genetic susceptibility to diclofenac-induced hepatotoxicity: contribution of UGT2B7, CYP2C8, and abacavir C2 genotypes. Gastroenterology 2007;132(1):272–81.

[125] Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe'er I, Floratos A, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. Nat Genet 2009;41(7):816–9.

[126] Spraggs CF, Budde LR, Briley LP, Bing N, Cox CJ, King KS, et al. HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. Int J Clin Oncol 2011;29(6):667–73.

[127] Batchelor JR, Welsh KI, Tinoco RM, Dollery CT, Hughes GR, Bernstein R, et al. Hydralazine-induced systemic lupus erythematosus: influence of HLA-DR and sex on susceptibility. Lancet 1980;1(8178):1107–9.

[128] Vedove CD, Del Giglio M, Schena D, Girolomoni G. Drug-induced lupus erythematosus. Arch Dermatol Res 2009;301(1):99–105.

[129] Kim SH, Choi JH, Lee KW, Kim SH, Shin ES, Oh HB, et al. The human leukocyte antigen-DRB1*1302-DQB1*0609-DPB1*0201 haplotype may be a strong genetic marker for aspirin-induced urticaria. Clin Exp Allergy 2005;35(3):339–44.

[130] Kim SH, Hur GY, Choi JH, Park HS. Pharmacogenetics of aspirin-intolerant asthma. Pharmacogenomics 2008;9(1):85–91.

[131] Rodriguez-Perez M, Gonzalez-Dominguez J, Mataran L, Garcia-Perez S, Salvatierra D. Association of HLA-DR5 with mucocutaneous lesions in patients with rheumatoid arthritis receiving gold sodium thiomalate. J Rheumatol 1994;21(1):41–3.

[132] Speerstra F, Reekers P, van de Putte LB, Vandenbroucke JP. HLA associations in aurothioglucose- and D-penicillamine-induced haematotoxic reactions in rheumatoid arthritis. Tissue Antigens 1985;26(1):35–40.

[133] Quiralte J, Sanchez-Garcia F, Torres MJ, Blanco C, Castillo R, Ortega N, et al. Association of HLA-DR11 with the anaphylactoid reaction caused by nonsteroidal anti-inflammatory drugs. J Allergy Clin Immunol 1999;103(4):685–9.

[134] Garlepp MI, Dawkins RL, Christiansen FT. HLA antigens and acetylcholine receptor antibodies in penicillamine induced myasthenia gravis. Br Med J 1983;286(6375):1442–3.

[135] Maier LA, McGrath DS, Sato H, Lympany P, Welsh K, Du Bois R, et al. Influence of MHC class II in susceptibility to beryllium sensitization and chronic beryllium disease. J Immunol 2003;171(12):6910–8.

[136] Rosenman KD, Rossman M, Hertzberg V, Reilly MJ, Rice C, Kanterakis E, et al. HLA class II DPB1 and DRB1 polymorphisms associated with genetic susceptibility to beryllium toxicity. Occup Environ Med 2011;68(7):487–93.

[137] Pellicano R, Ciavarella G, Lomuto M, Di Giorgio G. Genetic susceptibility to fixed drug eruption: evidence for a link with HLA-B22. J Am Acad Dermatol 1994;30(1):52–4.

[138] Pellicano R, Lomuto M, Ciavarella G, Di Giorgio G, Gasparini P. Fixed drug eruptions with feprazone are linked to HLA-B22. J Am Acad Dermatol 1997;36(5 Pt 1):782–4.

[139] Ozkaya-Bayazit E, Akar U. Fixed drug eruption induced by trimethoprim-sulfamethoxazole: evidence for a link to HLA-A30 B13 Cw6 haplotype. J Am Acad Dermatol 2001;45(5):712–7.

[140] Lieberman JA, Yunis J, Egea E, Canoso RT, Kane JM, Yunis EJ. HLA-B38, DR4, DQw3 and clozapine-induced agranulocytosis in Jewish patients with schizophrenia. Arch Gen Psychiatry 1990;47(10):945–8.

[141] Valevski A, Klein T, Gazit E, Meged S, Stein D, Elizur A, et al. HLA-B38 and clozapine-induced agranulocytosis in Israeli Jewish schizophrenic patients. Eur J Immunogenet 1998;25(1):11–13.

[142] Athanasiou MC, Dettling M, Cascorbi I, Mosyagin I, Salisbury BA, Pierz KA, et al. Candidate gene analysis identifies a polymorphism in HLA-DQB1 associated with clozapine-induced agranulocytosis. J Clin Psychiatry 2011;72(4):458–63.

[143] Dettling M, Cascorbi I, Opgen-Rhein C, Schaub R. Clozapine-induced agranulocytosis in schizophrenic Caucasians: confirming clues for associations with human leukocyte class I and II antigens. Pharmacogenomics J 2007;7(5):325–32.

[144] Diez RA. HLA-B27 and agranulocytosis by levamisole. Immunol Today 1990;11(8):270.

[145] Nolan D, Gaudieri S, Mallal S. Pharmacogenetics: a practical role in predicting antiretroviral drug toxicity? J HIV Ther 2003;8(2):36–41.

[146] Phillips E, Nolan, D, Thorborn, D, Schaefer, M, Laird, R, Rauch, A, et al., (Eds.), Genetic factors predicting abacavir hypersensitivity

and tolerance in HLA-B*5701 positive individuals. 3rd International Drug Hypersensitivity Conference; 2008; Paris, France. PMID: 18652038.

[147] Hughes AR, Mosteller M, Bansal AT, Davies K, Haneline SA, Lai EH, et al. Association of genetic variations in HLA-B region with hypersensitivity to abacavir in some, but not all, populations. Pharmacogenomics 2004;5(2):203–11.

[148] Knowles SR, Shapiro LE, Shear NH. Anticonvulsant hypersensitivity syndrome: incidence, prevention and management. Drug Saf 1999;21(6):489–501.

[149] Mehta TY, Prajapati LM, Mittal B, Joshi CG, Sheth JJ, Patel DB, et al. Association of HLA-B*1502 allele and carbamazepine-induced Stevens-Johnson syndrome among Indians. Indian J Dermatol Venereol Leprol 2009;75(6):579–82.

[150] Chen P, Lin JJ, Lu CS, Ong CT, Hsieh PF, Yang CC, et al. Carbamazepine-induced toxic effects and HLA-B*1502 screening in Taiwan. N Engl J Med 2011;364(12):1126–33.

[151] Ko TM, Chung WH, Wei CY, Shih HY, Chen JK, Lin CH, et al. Shared and restricted T-cell receptor use is crucial for carbamazepine-induced Stevens-Johnson syndrome. J Allergy Clin Immunol 2011;128(6):1266–76. e11

[152] Mizumoto K, Sumikawa Y, Niihara H, Morita E. Case of carbamazepine-induced hypersensitivity syndrome associated with human leukocyte antigen-A*3101. J Dermatol 2012;39(9):791–2.

[153] Yip VL, Marson AG, Jorgensen AL, Pirmohamed M, Alfirevic A. HLA genotype and carbamazepine-induced cutaneous adverse drug reactions: a systematic review. Clin Pharmacol Ther 2012;92(6):757–65.

[154] Amstutz U, Ross CJ, Castro-Pastrana LI, Rieder MJ, Shear NH, Hayden MR, et al. HLA-A*31:01 and HLA-B*15:02 as genetic markers for carbamazepine hypersensitivity in children. Clin Pharmacol Ther 2013;94(1):142–9.

[155] Shi YW, Min FL, Qin B, Zou X, Liu XR, Gao MM, et al. Association between HLA and Stevens-Johnson syndrome induced by carbamazepine in Southern Han Chinese: genetic markers besides B*1502? Basic Clin Pharmacol Toxicol 2012;111(1):58–64.

[156] Grover S, Kukreti R. HLA allelic variants and carbamazepine-induced hypersensitivity. Clin Pharmacol Ther 2013;93(5):386–7.

[157] Dainichi T, Uchi H, Moroi Y, Furue M. Stevens-Johnson syndrome, drug-induced hypersensitivity syndrome and toxic epidermal necrolysis caused by allopurinol in patients with a common HLA allele: what causes the diversity? Dermatology 2007;215(1):86–8.

[158] Jung JW, Song WJ, Kim YS, Joo KW, Lee KW, Kim SH, et al. HLA-B58 can help the clinical decision on starting allopurinol in patients with chronic renal insufficiency. Nephrol Dial Transplant 2011;26(11):3567–72.

[159] Gandhi G, Buttar BS, Albert L, Hasan Q, Aggarwal RK. Psoriasis-associated genetic polymorphism in North Indian population in the CCHCR1 gene and in a genomic segment flanking the HLA-C region. Dis Markers 2011;31(6):361–70.

[160] Quinones-Lombrana A, Lopez-Soto A, Ballina-Garcia FJ, Alperi-Lopez M, Queiro-Silva R, Lopez-Vazquez A, et al. BAT1 promoter polymorphism is associated with rheumatoid arthritis susceptibility. J Rheumatol 2008;35(5):741–4.

[161] Shibue T, Tsuchiya N, Komata T, Matsushita M, Shiota M, Ohashi J, et al. Tumor necrosis factor alpha 5′-flanking region, tumor necrosis factor receptor II, and HLA-DRB1 polymorphisms in Japanese patients with rheumatoid arthritis. Arthritis Rheum 2000;43(4):753–7.

[162] Pompeu YA, Stewart JD, Mallal S, Phillips E, Peters B, Ostrov DA. The structural basis of HLA-associated drug hypersensitivity syndromes. Immunol Rev 2012;250(1):158–66.

[163] Pollard RB, Robinson P, Dransfield K. Safety profile of nevirapine, a nonnucleoside reverse transcriptase inhibitor for the treatment of human immunodeficiency virus infection. Clin Ther 1998;20(6):1071–92.

[164] Gangar M, Arias G, O'Brien JG, Kemper CA. Frequency of cutaneous reactions on rechallenge with nevirapine and delavirdine. Ann Pharmacother 2000;34(7–8):839–42.

[165] Bersoff-Matcha SJ, Miller WC, Aberg JA, van Der Horst C, Hamrick Jr HJ, Powderly WG, et al. Sex differences in nevirapine rash. Clin Infect Dis 2001;32(1):124–9.

[166] Bonjoch A, Paredes R, Domingo P, Cervantes M, Pedrol E, Ribera E, et al. Long-term safety and efficacy of nevirapine-based approaches in HIV type 1-infected patients. AIDS Res Hum Retroviruses 2006;22(4):321–9.

[167] Gonzalez de Requena D, Nunez M, Jimenez-Nacher I, Soriano V. Liver toxicity caused by nevirapine. AIDS 2002;16(2):290–1.

[168] Peters PJ, Stringer J, McConnell MS, Kiarie J, Ratanasuwan W, Intalapaporn P, et al. Nevirapine-associated hepatotoxicity was not predicted by CD4 count >/=250 cells/muL among women in Zambia, Thailand and Kenya. HIV Med 2010;11(10):650–60.

[169] Torti C, Costarelli S, De Silvestri A, Quiros-Roldan E, Lapadula G, Cologni G, et al. Analysis of severe hepatic events associated with nevirapine-containing regimens: CD4+ T-cell count and gender in hepatitis C seropositive and seronegative patients. Drug Saf 2007;30(12):1161–9.

[170] Kesselring AM, Wit FW, Sabin CA, Lundgren JD, Gill MJ, Gatell JM, et al. Risk factors for treatment-limiting toxicities in patients starting nevirapine-containing antiretroviral therapy. AIDS 2009;23(13):1689–99.

[171] Wit FW, Kesselring AM, Gras L, Richter C, van der Ende ME, Brinkman K, et al. Discontinuation of nevirapine because of hypersensitivity reactions in patients with prior treatment experience, compared with treatment-naive patients: the ATHENA cohort study. Clin Infect Dis 2008;46(6):933–40.

[172] Phillips E, Bartlett JA, Sanne I, Lederman MM, Hinkle J, Rousseau F, et al. Associations between HLA-DRB1*0102, HLA-B*5801, and hepatotoxicity during initiation of nevirapine-containing regimens in South Africa. J Acquir Immune Defic Syndr 2013;62(2):e55–7.

[173] Badulli C, Sestini R, Sbarsi I, Baroncelli M, Pizzochero C, Martinetti M, et al. Tag SNPs of the ancestral haplotype 57.1 do not substitute HLA-B*57:01 typing for eligibility to abacavir treatment in the Italian population. Pharmacogenomics 2012;13(3):247–9.

[174] Andrade RJ, Lucena MI, Alonso A, Garcia-Cortes M, Garcia-Ruiz E, Benitez R, et al. HLA class II genotype influences the type of liver injury in drug-induced idiosyncratic liver disease. Hepatology 2004;39(6):1603–12.

[175] Lucena MI, Andrade RJ, Fernandez MC, Pachkoria K, Pelaez G, Duran JA, et al. Determinants of the clinical expression of amoxicillin-clavulanate hepatotoxicity: a prospective series from Spain. Hepatology 2006;44(4):850–6.

[176] Grant LM, Rockey DC. Drug-induced liver injury. Curr Opin Gastroenterol 2012;28(3):198–202.

[177] Andrade RJ, Lucena MI, Fernandez MC, Pelaez G, Pachkoria K, Garcia-Ruiz E, et al. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. Gastroenterology 2005;129(2):512–21.

[178] Stephens C, Lopez-Nevot MA, Ruiz-Cabello F, Ulzurrun E, Soriano G, Romero-Gomez M, et al. HLA alleles influence the clinical signature of amoxicillin-clavulanate hepatotoxicity. PLoS ONE 2013;8(7):e68111.

[179] Donaldson PT, Daly AK, Henderson J, Graham J, Pirmohamed M, Bernal W, et al. Human leukocyte antigen class II genotype in susceptibility and resistance to co-amoxiclav-induced liver injury. J Hepatol 2010;53(6):1049–53.

[180] Staszewski S, Keiser P, Montaner J, Raffi F, Gathe J, Brotas V, et al. Abacavir-lamivudine-zidovudine vs indinavir-lamivudine-zidovudine in antiretroviral-naive HIV-infected adults: a randomized equivalence trial. JAMA 2001;285(9):1155–63.

[181] DeJesus E, Herrera G, Teofilo E, Gerstoft J, Buendia CB, Brand JD, et al. Abacavir versus zidovudine combined with lamivudine and efavirenz, for the treatment of antiretroviral-naive HIV-infected adults. Clin Infect Dis 2004;39(7):1038–46.

[182] Gulick RM, Ribaudo HJ, Shikuma CM, Lalama C, Schackman BR, Meyer 3rd WA, et al. Three- vs four-drug antiretroviral regimens for the initial treatment of HIV-1 infection: a randomized controlled trial. JAMA 2006;296(7):769–81.

[183] Osawa J, Naito S, Aihara M, Kitamura K, Ikezawa Z, Nakajima H. Evaluation of skin test reactions in patients with non-immediate type drug eruptions. J Dermatol 1990;17(4):235–9.

[184] Alanko K. Patch testing in cutaneous reactions caused by carbamazepine. Contact dermatitis 1993;29(5):254–7.

[185] Schnyder B, Adam J, Rauch A, Thurnheer MC, Pichler WJ. HLA-B*57:01 abacavir-naive individuals have specific T cells but no patch test reactivity. J Allergy Clin Immunol 2013;132(3):756–8.

[186] Leckband SG, Kelsoe JR, Dunnenberger HM, George Jr AL, Tran E, Berger R, et al. Clinical pharmacogenetics implementation consortium guidelines for HLA-B genotype and carbamazepine dosing. Clin Pharmacol Ther 2013;94(3):324–8.

[187] Martin MA, Klein TE, Dong BJ, Pirmohamed M, Haas DW, Kroetz DL, et al. Clinical pharmacogenetics implementation consortium guidelines for HLA-B genotype and abacavir dosing. Clin Pharmacol Ther 2012;91(4):734–8.

[188] Martin AM, Nolan D, Gaudieri S, Almeida CA, Nolan R, James I, et al. Predisposition to abacavir hypersensitivity conferred by HLA-B*5701 and a haplotypic Hsp70-Hom variant. Proc Nat Acad Sci USA 2004;101(12):4180–5.

[189] Hammond E, Almeida CA, Mamotte C, Nolan D, Phillips E, Schollaardt TA, et al. External quality assessment of HLA-B*5701 reporting: an international multicentre survey. Antivir Ther 2007;12(7):1027–32.

[190] Hammond E, Mamotte C, Nolan D, Mallal S. HLA-B*5701 typing: evaluation of an allele-specific polymerase chain reaction melting assay. Tissue Antigens 2007;70(1):58–61.

[191] Sayer DC, Cassell HS, Christiansen FT. HLA-B*27 typing by sequence specific amplification without DNA extraction. Mol Pathol 1999;52(5):300–1.

[192] Badulli C, Sbarsi I, Di Giorgio D, Mantovani M, Maserati R, Barbarini G, et al. A new approach to safely type for HLA the HIV infected people eligible to abacavir therapy: saliva or buccal swab as reliable DNA sources. Clin Chim Acta 2011;412(21–22):1995–8.

[193] Martin AM, Krueger R, Almeida CA, Nolan D, Phillips E, Mallal S. A sensitive and rapid alternative to HLA typing as a genetic screening test for abacavir hypersensitivity syndrome. Pharmacogenet Genomics 2006;16(5):353–7.

[194] Kostenko L, Kjer-Nielsen L, Nicholson I, Hudson F, Lucas A, Foley B, et al. Rapid screening for the detection of HLA-B57 and HLA-B58 in prevention of drug hypersensitivity. Tissue Antigens 2011;78(1):11–20.

[195] Colombo S, Rauch A, Rotger M, Fellay J, Martinez R, Fux C, et al. The HCP5 single-nucleotide polymorphism: a simple screening tool for prediction of hypersensitivity reaction to abacavir. J Infect Dis 2008;198(6):864–7.

[196] Chui CK, Brumme ZL, Brumme CJ, Yip B, Phillips EJ, Montaner JS, et al. A simple screening approach to reduce B*5701-associated abacavir hypersensitivity on the basis of sequence variation in HIV reverse transcriptase. Clin Infect Dis 2007;44(11):1503–8.

[197] Waters LJ, Mackie NE, Pozniak A, Winston A. Polymorphisms at position 245 of HIV reverse transcriptase do not accurately predict the presence of human leukocyte antigen B*5701. Clin Infect Dis 2007;45(9):1243–4. author reply 4–5

[198] Gazzard BG, Anderson J, Babiker A, Boffito M, Brook G, Brough G, et al. British HIV association guidelines for the treatment of HIV-1-infected adults with antiretroviral therapy 2008. HIV Med 2008;9(8):563–608.

[199] Aberg JA, Kaplan JE, Libman H, Emmanuel P, Anderson JR, Stone VE, et al. Primary care guidelines for the management of persons infected with human immunodeficiency virus: 2009 update by the HIV medicine association of the infectious diseases society of America. Clin Infect Dis 2009;49(5):651–81.

[200] Becquemont L, Alfirevic A, Amstutz U, Brauch H, Jacqz-Aigrain E, Laurent-Puig P, et al. Practical recommendations for pharmacogenomics-based prescription: 2010 ESF-UB conference on pharmacogenetics and pharmacogenomics. Pharmacogenomics 2011;12(1):113–24.

[201] Swen JJ, Nijenhuis M, de Boer A, Grandia L, Maitland-van der Zee AH, Mulder H, et al. Pharmacogenetics: from bench to byte–an update of guidelines. Clin Pharmacol Ther 2011;89(5):662–73.

[202] Ferrell Jr PB, McLeod HL. Carbamazepine, HLA-B*1502 and risk of Stevens-Johnson syndrome and toxic epidermal necrolysis: US FDA recommendations. Pharmacogenomics 2008;9(10):1543–6.

[203] Hershfield MS, Callaghan JT, Tassaneeyakul W, Mushiroda T, Thorn CF, Klein TE, et al. clinical pharmacogenetics implementation consortium guidelines for human leukocyte antigen-B genotype and allopurinol dosing. Clin Pharmacol Ther 2013;93(2):153–8.

This page is intentionally left blank

Chapter 22

Drug-Induced Liver Injury

Thomas J. Urban[a] **and Hans Tillmann**[b]
[a]*Center for Human Genome Variation, USA,* [b]*Department of Medicine, Division of Gastroenterology,*
Duke University Medical Center, Durham, NC, USA

Chapter Outline

22.1 Overview 467
22.2 Epidemiology and Public Health Impact 467
22.3 Etiology 468
22.4 Pharmacogenomics and Stratified Medicine in DILI 469
  22.4.1 Potential Value of Understanding Causes of DILI 469
  22.4.2 Potential Value of Predictive Biomarkers 469
  22.4.3 Potential Value of DILI Diagnostics 470
22.5 Difficulties in Identifying Biomarkers for DILI 470
  22.5.1 Acquiring DILI Cases 470
  22.5.2 Diagnosis and Phenotyping 471
  22.5.3 DILI Heterogeneity 471
  22.5.4 Lack of Adequate Model Systems 471
22.6 DILI-Associated Genetic Variants 471
  22.6.1 DILI Genetics in the Candidate Gene Era 471
  22.6.2 Human Leukocyte Antigen Genes and DILI Risk 472
  22.6.3 Functional Characterization of DILI-Associated Alleles 472
22.7 Utility of Dili Risk Variants in the Clinic 472
22.8 Future Directions 473
  22.8.1 Expanded-Content Genotyping Arrays 473
  22.8.2 Next-Generation Sequencing 474
  22.8.3 Epigenomics 474
  22.8.4 Induced Pluripotent Stem Cells 474
22.9 Conclusions 475
References 475

22.1 OVERVIEW

Despite advances in the safety of therapeutic drugs over the past half century, adverse drug reactions (ADRs) remain a major problem in healthcare. ADRs are estimated to be the fourth leading cause of death in the United States [1], associated with more than 100,000 deaths annually, and they are responsible for approximately 7% of hospital admissions [2]. The financial burden associated with ADRs is staggering, with estimates as high as \$177 billion in treatment costs per year, representing up to 10% of gross health expenditures in the United States [3,4].

22.2 EPIDEMIOLOGY AND PUBLIC HEALTH IMPACT

Among ADRs, rare, serious drug-induced adverse events (SAEs) pose significant problems for patients, healthcare providers, payers, and regulators, and are an important element contributing to healthcare cost inflation. Examples of SAEs include hepatotoxicity, severe QT prolongation, rhabdomyolysis, serious skin rashes (e.g., Stevens-Johnson syndrome), acute renal failure, acute hypersensitivity, anemias/neutropenias, visual field loss, and vasculitis, among others. Although such severe adverse events are rare, their severity nonetheless often leads to reduced utilization of the associated drug(s), or their withdrawal from the market. Between 1976 and 2005, 28 drugs were withdrawn from the market, primarily due to rare hepatic, renal, or cardiovascular toxicity that was not anticipated from animal toxicity studies or from phase I–III clinical trials [5]. Thus, in addition to clinical costs, the financial losses caused by SAEs identified after market approval are compounded by the high and rising cost of drug development, currently estimated to be more than \$1.5 billion per approved drug.

Our focus in this chapter is on one of the most important categories of drug-induced severe adverse events. Severe hepatotoxicity, or drug-induced liver injury (DILI), is the most common reason for post-approval drug failures [5]. DILI, excluding acetaminophen associated injury, is estimated to be responsible for approximately 11% of all cases

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00022-0
© 2014 Elsevier Inc. All rights reserved.

of acute liver failure [6], with an estimated 200 DILI cases in the United States annually [5].

It is unclear how many different drugs are capable of causing DILI, but the most recent estimates are at least in the hundreds [7,8]. Among these, there are several that are the most highly represented: amoxicillin/clavulanate, diclofenac, flucloxacillin, isoniazid, nimesulide, sulfamethoxazole/trimethoprim, tetracyclines, and valproic acid. The fact that these are among the most frequent causes of DILI does not necessarily imply that they are particularly toxic, but it may be that they are very frequently prescribed. Overall estimates of the incidence of DILI range as high as 13.9 per 100,000 patients per year [9], but varies widely depending on the drug. For example, DILI due to valproic acid has been estimated to occur in as many as 1 in 500 patients in certain high-risk groups [10].

22.3 ETIOLOGY

The vast majority of DILI is due to idiosyncratic reactions not predictable from drug dosage/concentration [11], which has led to the suspicion that at least some of the risk for DILI is due to intrinsic genetic risk factors, some of which have already been identified. Many different drugs can cause DILI, with the precise pattern of injury varying among them, and even among patients with liver injury due to the same drug.

Typically, DILI reactions are classified as "hepatocellular" when the injury is believed to be focused on the hepatocyte and "cholestatic" when the damage is believed to occur at the hepatocyte canalicular membrane or further downstream in the biliary tree [12]. The precise mechanisms of either type of reaction are unclear; however, there is emerging evidence that both direct effects of the drug or its metabolites, and immune responses to injury, play a role.

Figure 22.1 shows a general model for how drugs might cause liver toxicity. As discussed in more detail below, it is thought that drug metabolism, or "bioactivation," may be part of the etiology of DILI [13]. Likewise, the ability of drugs to inhibit bile acid efflux from hepatocytes, either at the bile canalicular membrane or at the basolateral membrane, may contribute to bile acid congestion in the liver and lead to liver injury [14]. For at least some drugs, there is evidence that mitochondrial toxicity may be a contributing factor for DILI, perhaps by reducing the rate of hepatocyte regeneration [15]. Finally, there is considerable evidence from human genetic studies that adaptive immune mechanisms contribute to DILI, with both antigen presentation by the liver (through HLA Class I proteins) and recognition of antigen by professional antigen-presenting cells (APCs) implicated in the development of DILI for at least some drugs [16].

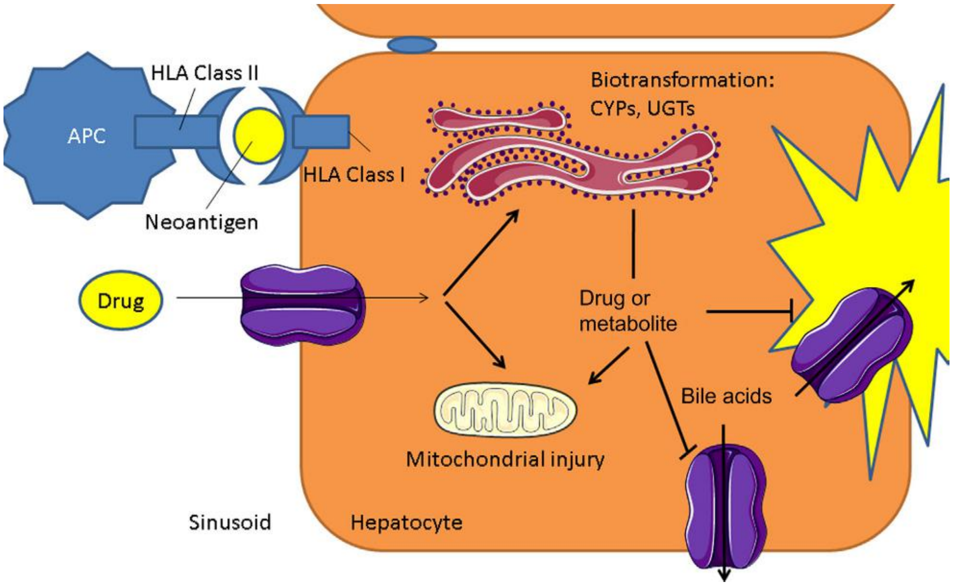


FIGURE 22.1 **Mechanisms for DILI.** It is currently hypothesized that drugs cause stress to hepatocytes, possibly through generation of reactive metabolites. Such stresses may include mitochondrial injury or the accumulation of bile acids in the hepatocyte due to inhibition of bile acid efflux. The magnitude of the stress to the hepatocyte may reflect the balance between production and elimination of the toxic metabolite(s), as well as the ability of the liver to respond or regenerate in the face of injury. Although such stress could lead to hepatocyte death by itself, there is growing evidence that death may involve, or at least be propagated by, both innate and adaptive immune responses, with particular involvement of human leukocyte antigen (HLA) molecules. It is believed that the role for HLA in the development of DILI involves (1) presentation of the neoantigen (possibly the parent drug, a metabolite, or a drug-protein conjugate) on the surface of the hepatocyte by HLA Class I molecules and (2) stimulation of T-cell responses by HLA Class II presentation on antigen-presenting cells (APCs).

22.4 PHARMACOGENOMICS AND STRATIFIED MEDICINE IN DILI

Given the substantial pharmacoeconomic impact of liver toxicity at the drug development stage, this is perhaps the area in which improvements in our knowledge of the mechanisms and genetic determinants of DILI will have the highest value. If there were tools available to reliably screen drug candidates for potential liver toxicity at the preclinical or early clinical development phases, it might be possible to filter out compounds with the potential to cause hepatotoxicity so that the need for clinical predictors would not exist.

22.4.1 Potential Value of Understanding Causes of DILI

Testing for hepatotoxicity in preclinical animal models, particularly in rats, is currently a standard part of toxicity screening in drug development; however, the fact is that many drugs that show hepatotoxicity in humans are not identified as hepatotoxic in rats or other preclinical models, suggesting that the potential for hepatotoxicity is species-dependent or dependent on individual genetic characteristics that are not well represented in animal models [17]. Similarly, the use of primary hepatocytes or hepatocyte-derived cell lines (including HepG2 and more recently HepRG) does not appear to provide a reliably sensitive screen for *in vivo* hepatotoxicity in humans. The reasons for this are numerous, but can be distilled into a single explanation: our current knowledge of the mechanisms for hepatotoxicity, including the contribution of different types of immune cells and their interplay with compromised hepatocytes, is quite incomplete, and thus so is our ability to accurately model this process in cell culture. A better understanding of the mechanisms of hepatotoxicity in humans will allow us to build the appropriate cell culture systems to effectively capture the process and produce a reliable screening tool for liver toxicity.

A prototypical example of this model has recently become a reality for a different type of rare, severe drug toxicity: drug-induced cardiac arrhythmias, including torsades de pointes (TdP). It has been noted that drugs capable of causing TdP tend to prolong the QT interval by inhibition of the human Ether-a-go-go–related protein (hERG), an inward-rectifying potassium channel subunit. Thus, using cell culture models in which the inhibition of hERG can be detected by electrophysiology, for example, provides a highly sensitive screen for drugs with arrhythmogenic potential [18]. Although such methods are not perfect, and must be augmented with early-phase clinical studies that directly measure a drug's effect on QT interval prolongation, they have contributed significantly to reducing the number of arrhythmogenic drugs brought to late-phase clinical trials or to the market [18,19]. Now that liver toxicity appears to have become the major reason for post-approval drug withdrawal, a similarly effective screening method would be tremendously valuable to the pharmaceutical industry, and to patients and health care providers.

22.4.2 Potential Value of Predictive Biomarkers

While eliminating the risk for DILI from newly developed drugs would be a major advance in healthcare, the fact remains that a large number of already approved drugs have the potential to cause DILI and yet are widely used clinically—and are likely to continue to be used routinely until and unless they become supplanted by safer and less expensive alternatives. For this reason, identification of predictive biomarkers (e.g., genetic markers) for DILI would provide a much needed tool for clinicians to test patients prior to prescribing a drug in order to prevent DILI in susceptible patients. Also, without a biomarker, DILI currently is a diagnosis of exclusion.

What would an ideal DILI biomarker look like? For a predictive marker for DILI to be useful in the clinic, it must have certain characteristics, as described in the following subsections.

22.4.2.1 The Biomarker Does Not Change Over Time

For DILI, as for other SAEs, significant investment has been made to identify heritable genetic predictors of susceptibility. There are several reasons that genetic studies of DILI have been favored over other types of research, such as epidemiological studies of environmental DILI risk factors or studies of the role of gene expression in DILI risk, but a major reason is that, unlike environmental triggers or gene expression profiles, genomic DNA sequence variation remains unchanged throughout an individual's lifespan. This means that a reliable genetic test for DILI susceptibility could be performed at birth, or at any point in a patient's medical history, and remain useful for as long as such a test is deemed clinically relevant.

22.4.2.2 The Biomarker is Relevant for a Variety of Drugs

Ideally, a predictive test for DILI susceptibility would indicate whether a particular patient is generally at risk for DILI, as opposed to being at risk for particular causes or types of DILI, or at risk for DILI only in particular environments. Based on the results of genetic studies performed to date, it appears more likely that risk factors for DILI more often than not act in a drug-specific manner [20], or at least are relevant to only a small number of hepatotoxic drugs. Still, there can only be a finite number of biological causes of DILI, and if most of the risk for it can be partitioned into a small number of biological pathways and genetic susceptibility alleles, it may be possible for a "DILI chip" or panel of biomarkers to be developed to give a broad estimate of DILI susceptibility, including drug-specific risk evaluation for an individual patient. The current state of the science suggests that this will not be possible using common SNPs [7], but this does not imply that such markers will not be found using more contemporary methods such as whole-exome or whole-genome sequencing.

22.4.2.3 The Negative Predictive Value is Very High (Especially for Drugs that Cause DILI Relatively Frequently)

One way in which predictive markers for DILI may be useful is in reducing its incidence by finding a common feature among DILI patients that occurs at a reasonably low frequency in the general population. Even if this means depriving a large number of patients of access to a particular drug, if serious toxicity can be minimized or even eliminated, this may still be a useful tool for prescribers, at least in certain contexts. For example, it was shown that in patients of Han Chinese ancestry who developed Stevens-Johnson syndrome (SJS) due to carbamazepine, the carrier frequency of HLA-B*1502 was 100%, compared with only 3% in carbamazepine-tolerant patients and approximately 9% in the general population [21]. Thus, HLA-B*1502 carrier status was shown to be a perfectly sensitive marker for carbamazepine-induced SJS. More recently, this genetic marker was evaluated in a prospective, genotype-guided trial, in which it was shown that SJS could be entirely prevented by genotyping for HLA-B*1502 prior to prescribing carbamazepine [22]. This test is now routinely used in the clinic and is standard practice prior to carbamazepine prescriptions in at least 50 hospitals in Taiwan.

22.4.2.4 The Positive Predictive Value is Very High (Especially for Drugs that Rarely Cause DILI)

Of course, optimally, a predictive test for DILI would tell us exactly which patients are very likely to experience it, especially when the drug being considered is widely prescribed and generally safe (i.e., the incidence of DILI from the drug is very rare). In such cases, a common SNP marker or HLA type may be strongly associated with DILI but not be a clinically useful predictor, merely because the vast majority of patients who are positive for the marker would not be at risk for DILI and would therefore be denied a useful medicine for no real benefit. On the other hand, a rare marker (e.g., a rare genetic variant) that was perfectly correlated with DILI would be rather useful. To date, the study of rare genetic variants in the context of DILI has been limited, but genomic technology has now advanced to the point where such predictors may soon be found, and will be of immediate clinical utility [20].

22.4.2.5 The Biomarker can be Measured Using Easily Accessible Patient Biological Material

Unlike pharmacogenetic/pharmacogenomic studies in cancer, where abundant material from the affected tissue (i.e., the tumor) is often available from diagnostic biopsies or surgical excision of cancerous tissue, it is unusual for clinicians to have access to liver tissue in advance of a clinical indication for biopsy; thus, any predictive test for DILI will almost certainly need to be performed using more accessible tissue, such as blood or saliva, as opposed to "diseased" tissue taken by biopsy or under surgery.

22.4.3 Potential Value of DILI Diagnostics

Given that DILI is currently considered a "diagnosis of exclusion"—in other words, that attributing a particular case of liver injury to a drug primarily relies on ruling out competing causes (e.g., viral hepatitis, autoimmune disease)—it would be of some benefit to clinicians if there were a reliable test that could be given after the occurrence of suspected DILI that would increase confidence in the diagnosis, even if such a test were not sufficiently informative to be used as a predictive marker. As discussed below, one such marker may already exist, with HLA-B*5701 carrier status being a potentially useful diagnostic for DILI due to flucloxacillin [23].

22.5 DIFFICULTIES IN IDENTIFYING BIOMARKERS FOR DILI

Investigation of the risk factors for DILI in humans, whether inherited or acquired, is limited primarily by the available sample sizes of cohorts of patients who have experienced it. Given that DILI is a very rare consequence of most drugs, ascertainment of cases requires a significant amount of diligence and investment.

22.5.1 Acquiring DILI Cases

Over the past decade, a number of national and international registries of DILI patients have been compiled. The US-based Drug Induced Liver Injury Network (DILIN), sponsored by the National Institutes of Health, has been active since 2003 and has enrolled more than 1,400 patients with DILI for the purposes of research into its causes and consequences [8,24]. Several European cohorts of DILI patients have similarly accrued large numbers of cases for research purposes, including DILIGEN [23], EUDRAGENE [25], and the Spanish DILI Registry [26]. Under the auspices of the International Serious Adverse Events Consortium (iSAEC, www.saeconsortium.org/), many of these independent efforts have been coordinated into large, collaborative genetic studies, and have led to a number of significant findings yielding insight into the genetic basis of DILI susceptibility, and continuing efforts promise to provide further discoveries in this area [23,27].

22.5.2 Diagnosis and Phenotyping

The study of drug-induced liver injury is complicated by the lack of specific tests to diagnose it, unlike, for example, viral hepatitis. Thus, causality assessment methods are needed. The most widely used method for causality assessment in DILI is the Roussel Uclaf Causality Assessment Method (RUCAM) [28,29], which is based on a series of objective questions designed to determine the causal relationship between drug initiation and liver injury. Recent efforts have improved on the RUCAM by using an expert panel-based assessment model [30,31]. These methods have likely improved the quality of DILI phenotyping in research studies; however, their implementation may not be practical in all cases, particularly in the clinic. Therefore, it may be that the gold standard for causality assessment includes consensus among experts on the diagnosis, but more practical methods must also be used. There have recently been attempts to standardize definitions and diagnostic approaches that may soon be established as general guidelines for the field [32].

22.5.3 DILI Heterogeneity

In addition to the problem of ambiguous diagnosis in the clinic, there is clearly also a very large degree of heterogeneity in the phenotype. Differences exist among DILI patients in terms of (1) the drug or drugs thought to be responsible for the injury, (2) the clinical presentation (including, for example, liver chemistries, delay between drug initiation and onset of injury, clinical signs and symptoms, and duration of injury), and (3) outcomes from liver injury, ranging from asymptomatic abnormal liver chemistries to reactions requiring liver transplantation or resulting in death [8,26,33].

It seems likely that this variation in DILI presentation from patient to patient may reflect intrinsic differences in the etiology of the injury. As evidence for this, and as described in more detail below, most of the successful genetic studies of DILI to date have been possible because of reasonably large sample sizes of DILI patients whose injury was due to the same drug or drug product, suggesting that at least genetic risk factors for DILI may tend to be drug-specific. On the other hand, it seems reasonable to expect that at least some of the risk factors for DILI are shared across different causal drugs. Thus, if DILI patients could be grouped into "clusters" based on structural similarity among causal drugs, similarities in clinical presentation, or common biological features, this might improve the prospects for discovery of both genetic and nongenetic risk factors by allowing for larger sample sizes of patients falling into these categories, relative to what is available for studies of individual DILI-causing drugs. Currently, however, such information on common mechanisms of DILI remains elusive.

22.5.4 Lack of Adequate Model Systems

As described above, good model systems of DILI, whether *in vitro* or animal models, are currently lacking, and this has significantly limited our ability to understand or to predict DILI. In addition, access to relevant tissues from human subjects is a major difficulty.

22.6 DILI-ASSOCIATED GENETIC VARIANTS

22.6.1 DILI Genetics in the Candidate Gene Era

For much of the history of DILI pharmacogenetics, there was a strong focus on genes or proteins involved in drug metabolism and disposition, with the hypothesis that interindividual variation in the pharmacokinetics of drugs might explain the variable risk of DILI among patients. Although the hypothesis was and remains compelling, the fact is that the recent successful genome-wide association studies (GWAS) have largely ruled out a major role for common variation in so-called "pharmacokinetic genes" in DILI risk. Still, there are many reports of positive associations between pharmacokinetic gene variants and liver injury, and while most of these remain unproven (in the sense of achieving contemporary standards of evidence in human genetic studies), there may be some utility in further studying these genes for their potential role in DILI pathogenesis.

The N-acetyl transferase gene NAT2 has been the most well studied gene in terms of its role in DILI, particularly hepatotoxicity due to isoniazid, which is metabolized by it. Although the results of the many studies on NAT2's role in isoniazid-induced liver injury are conflicting, the consensus is that NAT2 slow acetylators are at increased risk for liver toxicity from isoniazid [34]. It should be noted, however, that most of studies of NAT2 in isoniazid DILI have been performed in East Asian populations. A meta-analysis of NAT2 association studies showed that while NAT2 slow acetylation appears overall to increase risk for isoniazid toxicity, this has not been demonstrated in Caucasian patients [35]. The reasons for this lack of generalizability of the NAT2 results are unknown, but may be related to differences in dosing and monitoring practices in different parts of the world.

Several cytochrome p450 (CYP) genes have been associated with various types of DILI. Most notably, the CYP2E1 gene has been repeatedly implicated in liver toxicity due to isoniazid, with the CYP2E1*5 allele [36,37]. However, this allele was not associated with risk for isoniazid-induced liver injury in other study populations [38,39], and a recent meta-analysis of CYP2E1 studies does not support a role for genetic variation in this gene in isoniazid toxicity [35]. The CYP2C9 gene has been investigated for its potential role in liver injury due to diclofenac, and other drugs, with the results generally not supporting a significant role for genetic CYP2C9 variants [40,41].

The uridinyl diphosphate glucotransferases (UGTs) have also been studied for a role in liver injury, and evidence has been found both for a protective [42] and a deleterious [43] role for UGTs in the development of DILI. However, these results have come from relatively small studies of candidate genes and have not been replicated independently. Similar findings and conclusions have been found for glutathione S-transferase genes (GSTs), including GSTM1 and GSTT1, with the null alleles of these genes being considered risk alleles for DILI due to multiple drugs [44,45]. Again, these results have not been replicated and should be interpreted with caution.

In addition to drug-metabolizing enzymes, there has been a suggestion that certain membrane transporters may be involved in the pathology of DILI. These include genes such as ABCB4 (encoding MDR3), ABCB11 (encoding BSEP), and ABCC2 (encoding MRP2). These transporters are of particular interest in the study of DILI because they are capable of transporting bile acids in addition to their roles as multispecific drug transporters. Since bile acid accumulation in hepatocytes is thought to be one mechanism for the development of DILI, it seems reasonable to suspect that genetic variation in these transporters might be involved either because of their effects on drug disposition or because of the effects of drugs on bile acid handling in the liver. While a number of positive reports exist of genetic variants in these transporter genes relating to DILI risk [43,46,47], none have been confirmed in independent studies [7,48].

22.6.2 Human Leukocyte Antigen Genes and DILI Risk

In addition to genes involved in pharmacokinetics, immune-related genes such as those encoding the human leukocyte antigen (HLA) proteins and several cytokines have been nominally associated with risk for DILI even before the advent of genome-wide association studies (GWAS) [49]. That these genes were considered strong candidates for involvement in DILI reflects a long-standing view that DILI may be at least partly viewed as an immune mediated disease, though perhaps not appreciated until recently was the relative importance of variation in immune-related genes in DILI risk versus genes involved in drug metabolism and disposition.

Perhaps the best example of genetic determinants underlying susceptibility is DILI due to flucloxacillin, in which whole genome association (WGA) has identified the HLA-B*5701 allele as a major determinant of susceptibility. Patients with this allele have an 80-fold increase in susceptibility to DILI when treated with flucloxacillin [23]. Several common variants in the MHC region have been associated with this increase, which is caused by the antibiotic combination amoxicillin/clavulanate; in this case, both HLA Class I and HLA Class II genes are thought to contribute, and possibly interact with one another to modulate DILI risk [27]. Other HLA associations have been observed for lapatinib [50], lumiracoxib [51], and ximelagatran [52]. While it is clear that certain HLA alleles contribute to risk for DILI, it also appears that many of the risk alleles are drug-specific or at least do not contribute generally to this risk [7]. A list of genes in which polymorphisms have been associated with DILI is shown in Table 22.1.

22.6.3 Functional Characterization of DILI-Associated Alleles

While the role of HLA genes in DILI risk has been known or suspected for some time, it has been unclear until very recently how precisely particular HLA alleles cause its development. It has been hypothesized that the drug itself, a drug metabolite, or a protein-drug adduct (hapten) might act as a "neoantigen" and trigger an immune response. Recently, it was shown that carriers of the HLA-B*5701 allele, which has been associated with risk of DILI due to flucloxacillin, have activated CD8+ T cell responses when flucloxacillin is presented by dendritic cells [53]. Further, flucloxacillin was found to bind covalently to albumin, and the degree of binding was positively correlated with T-cell activation, suggesting that flucloxacillin-albumin adducts may be the neoantigen responsible for immune stimulation. This is in contrast to similar studies of the abacavir–HLA-B*5701 interaction [54]. These results give some insight into the processes leading to immune- or HLA-mediated DILI; however, much more work in this area is necessary to give a more complete understanding of the interplay between the liver and the immune system in the DILI's pathophysiology.

22.7 UTILITY OF DILI RISK VARIANTS IN THE CLINIC

Despite the fact that there are now several unequivocal associations between particular genetic variants and risk for DILI, currently there are no predictive tests that are being used clinically. The reasons for this include lack of positive predictive value, a high number needed to treat (NNT), incomplete negative predictive value, and the existence of therapeutic alternatives to drugs that frequently cause DILI.

TABLE 22.1 Genetic Polymorphisms Associated with DILI

Gene	Drug	Type of Study	References
HLA-B*5701	Flucloxacillin	GWAS	[23]
*HLA-DRB1*1501*	Amoxicillin-clavulanate	Candidate gene + GWAS	[27, 63, 51]
	Lumiracoxib	GWAS	
*HLA-DRB1*0701*	Ximelagatran	GWAS	[52, 50]
	Lapatinib	GWAS	
ABCC2 (MRP2)	Multiple drugs	Candidate gene	[43, 46]
ABCB11 (BSEP)	Multiple drugs	Candidate gene	[47]
CYP2E1	Isoniazid	Candidate gene	[36, 37]
GSTM1 and GSTT1 deletions	Multiple drugs	Candidate gene	[44]
NAT2	Isoniazid	Candidate gene	Reviewed in [34, 35]
POLG	Valproic acid	Candidate gene	[64]
SOD2	Multiple	Candidate gene	[65]
UGT1A6	Tolcapone	Candidate gene	[42]
UGT2B7	Diclofenac	Candidate gene	[43]

For most drugs that are currently in clinical use, the rate of DILI is extremely low (less than 1 in 10,000 patients). Thus, if only common alleles (whether SNPs, or HLA types) are known risk factors, then, by definition, most of the individuals carrying the risk allele are not truly susceptible to liver injury. The flucloxacillin-HLA-B*5701 example is once again informative: if all patients being considered for treatment with flucloxacillin were to undergo testing for HLA-B*5701 prior to treatment, it would take more than 10,000 genetic tests to prevent one case of DILI; in the meantime, 7% of patients who are carriers of HLA-B*5701, most of whom are not at risk, would be denied the first drug of choice [55]. In contrast, for a drug like ximelagatran, which causes DILI in up to 8% of patients, knowledge of HLA-DRB1*0701 carrier status might plausibly be used to prevent it: in this case, the NNT (or the number of patients to be genotyped in order to prevent one case of DILI) is a more reasonable 24, and approximately one-third of patients would be denied the drug because they are truly susceptible [55]. When evaluating a potential genetic test for the prevention of DILI, Such considerations of predictive value and NNT of course need to be considered in the clinical environment and in regard to possible therapeutic alternatives. In the case of ximelagatran, its potential benefits over existing anticoagulants might prove to be a reason to reconsider its clinical use if liver toxicity can be prevented by genetic testing.

While the predictive test for DILI remains elusive, several of the HLA associations with certain of its causes have been shown to have very high sensitivity and may in some cases have possible value as tests to aid in diagnosis after a suspected DILI event has occurred. For example, the HLA-B*5701 allele is carried by approximately 7% of individuals of European ancestry, but in roughly 85% of individuals who experienced flucloxacillin-induced liver injury [27]. Thus, in a case of suspected flucloxacillin-DILI, it might be reasonable to test for HLA-B*5701 carrier status: if the patient is found to be HLA-B*5701 positive, this might strengthen the case for a DILI diagnosis. However, since the population frequency of this allele is 7%, and most carriers are not truly at risk, this means that such a test will have roughly a 7% false positive rate—that is, HLA-B*5701 positivity may be entirely coincidental.

22.8 FUTURE DIRECTIONS

While GWASs have so far provided only limited information on the genetic causes of DILI, the continuing rapid advances in genomics and other technologies promise to expand the range of inquiry into its biological determinants of, and they may soon yield additional insight into DILI mechanisms and potentially clinically useful predictive biomarkers.

22.8.1 Expanded-Content Genotyping Arrays

Currently, there is much interest in expanding the GWAS concept, or at least the technology, to include not just a collection of common tagging SNP, but a much larger catalog of human genomic variation including rare variants (i.e., variants with population frequencies less than 1%). In the fall of 2012, several genotyping array manufacturers, in collaboration with a number of human geneticists in academia, developed genotyping platforms intended to cover a large fraction of protein-coding or exonic genetic variation, including variants with frequencies as low as 0.02% in the population (see http://genome.sph.umich.edu/wiki/Exome_Chip_Design). These so-called "exome chips" are currently a much more affordable and efficient approach to studying rare functional genetic variation on a large scale [56–58], and are currently being used to study a broad range of complex human traits, including severe adverse drug reactions (e.g., DILI).

Although their "coverage" of the protein-coding region of the genome is substantial, these arrays should not be considered a substitute for whole-genome or whole-exome sequencing for discovery of DILI risk factors, as any variants that are specific to DILI cases are, by definition, preferentially missing from these chips, whose content is based on genetic variation that is detectable in the general population. Still, it is certainly possible that protein-coding variants in the frequency range of 0.02–1%, which have not previously been interrogated for a role in DILI risk, may prove to contain many useful DILI biomarkers.

22.8.2 Next-Generation Sequencing

For essentially all of the known examples of genetic risk factors for DILI, the penetrance of the associated allele is very low; that is, the vast majority of individuals carrying the risk allele do not experience DILI after treatment with these drugs. Thus, it appears likely that the common risk variants require some modifier (e.g., an environmental trigger) or a rare genetic variant in an unlinked gene to produce DILI.

The advent of massively parallel or "next-generation" sequencing has revolutionized genetic discovery over the past several years. Most notable is the phenomenal pace of discovery of genetic causes of rare Mendelian diseases, with many previously unexplained diseases having been tracked to their genetic cause, often with only a handful of affected patients needed in order to securely implicate a particular gene in the etiology of disease [59]. Given that DILI is typically a rare event, with some drugs resulting in DILI in fewer than 1 in 10,000 patients per prescription, it is possible that, just like many rare diseases, DILI susceptibility may in fact be a Mendelian, monogenic trait. However, even if this were true, it would be unlikely that it would reveal itself clinically: Mendelian diseases are typically recognized on the basis of their segregating into families according to simple Mendelian rules of inheritance, but this is not likely be the case for DILI since it is very rare for entire families to be prescribed the same drug at the same dose concurrently. Thus, such Mendelian patterns of inheritance would be unlikely to emerge. Still, the hope that at least some causes of DILI will turn out to be "cryptic Mendelian traits" is currently fueling efforts to identify rare genetic causes of DILI using whole-exome and/or whole-genome sequencing [20].

One additional advantage to the growing use of whole-exome and whole-genome sequencing in the clinic is that we may soon reach the point at which most, or even all, patients are sequenced routinely and have their genomic sequence available on demand as part of their medical records. Once this happens, there will be many fewer arguments over the costs of predictive genetic testing, or the time between ordering a test and making prescribing decisions, because the test results will be sitting in the medical record to be inspected for any gene and any drug or indication. In such a scenario, the "number needed to treat" will become a largely irrelevant factor in the decision to use genetic information prior to prescribing a particular drug.

22.8.3 Epigenomics

Epigenetic DNA modification, including methylation at CpG dinucleotides, is a major mode of regulating gene expression. Recent technological advances in highly multiplex methylation arrays have allowed for epigenome-wide association studies [60], and have implicated DNA methylation in the pathogenesis of a host of common human diseases, including liver diseases such as nonalcoholic fatty liver disease (NAFLD) [61]. The role of DNA methylation, and other types of DNA modification, in the risk for DILI is as yet unexplored. One reason for this is, as noted above, the particular problem of accessing the relevant tissue because liver biopsies can rarely be obtained for research purposes from DILI patients. It is possible that epigenomic variation in circulating immune cells is responsible in part for DILI susceptibility, so that studies aimed at characterizing the epigenomic profile of peripheral blood leukocytes in DILI patients may reveal risk factors that have so far proven elusive. Additionally, as will be discussed, advances in stem cell technology may soon allow for the faithful reproduction of liver-like cell DILI culture models, which may allow for the use of epigenomic and other "omic" technologies in the absence of primary cells from the liver of DILI patients.

22.8.4 Induced Pluripotent Stem Cells

There has recently been considerable progress with culturing hepatocyte-like cells derived from induced pluripotent stem cells (iPSCs) [62]. In the near future, it is hoped that we will have the capacity to generate faithful cell culture models of DILI that can be used as models for understanding both the pathophysiology of DILI and the role of individual gene variants in DILI pathogenesis, and for screening compounds as potential hepatotoxins. Such models may include coculturing of hepatocytes with immune cells, and perhaps even three-dimensional models of liver incorporating immune cells, so that the relevant biology may be reproducible in the laboratory even without animal models. While this type of work might be done using cells from healthy donors, it may be that the most faithful models will be based on cells taken from those rare individuals who have experienced DILI. Such techniques will provide an immense resource for studying the biology of DILI and preventing the development of hepatotoxic drugs.

22.9 CONCLUSIONS

Drug-induced liver injury remains a major problem in clinical pharmacotherapy and drug development. The accumulating evidence for the role of adaptive immunity in the risk for DILI has provided some insight into the mechanisms underlying DILI, but to date there are no widely-accepted predictive biomarkers for DILI susceptibility. It is hoped that current and future studies that capitalize on the continuing advances in genomic technologies may further aid our understanding of the molecular causes of DILI and provide clinically useful pharmacogenomic tests for DILI risk.

REFERENCES

[1] Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;279(15):1200–5.
[2] Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 2004;329(7456):15–19.
[3] Ernst FR, Grizzle AJ. Drug-related morbidity and mortality: updating the cost-of-illness model. J Am Pharm Assoc 2001;41(2):192–9.
[4] Johnson JA, Bootman JL. Drug-related morbidity and mortality. A cost-of-illness model. Arch Intern Med 1995;155(18):1949–56.
[5] Wilke RA, Lin DW, Roden DM, Watkins PB, Flockhart D, Zineh I, et al. Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. Nat Rev Drug Discov 2007;6(11):904–16.
[6] Ostapowicz G, Fontana RJ, Schiodt FV, Larson A, Davern TJ, Han SH, et al. Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. Ann Intern Med 2002;137(12):947–54.
[7] Urban TJ, Shen Y, Stolz A, Chalasani N, Fontana RJ, Rochon J, et al. Limited contribution of common genetic variants to risk for liver injury due to a variety of drugs. Pharmacogenet Genomics 2012;22(11):784–95.
[8] Chalasani N, Fontana RJ, Bonkovsky HL, Watkins PB, Davern T, Serrano J, et al. Causes, clinical features, and outcomes from a prospective study of drug-induced liver injury in the United States. Gastroenterology 2008;135(6):1924–34. 34 e1–4.
[9] Sgro C, Clinard F, Ouazir K, Chanay H, Allard C, Guilleminet C, et al. Incidence of drug-induced hepatic injuries: a French population-based study. Hepatology 2002;36(2):451–5.
[10] Cotariu D, Zaidman JL. Valproic acid and the liver. Clin Chem 1988;34(5):890–7.
[11] Watkins PB. Idiosyncratic liver injury: challenges and approaches. Toxicol Pathol 2005;33(1):1–5.
[12] Council for International Organizations of Medical Sciences. Standardization of definitions and criteria of causality assessment of adverse drug reactions. Drug-induced liver disorders: report of an international consensus meeting. Int J Clin Pharmacol Ther 1990;28(8):317–22.
[13] Corsini A, Bortolini M. Drug-induced liver injury: the role of drug metabolism and transport. J Clin Pharmacol 2013;53(5):463–74.
[14] Yang K, Kock K, Sedykh A, Tropsha A, Brouwer KL. An updated review on drug-induced cholestasis: mechanisms and investigation of physicochemical properties and pharmacokinetic parameters. J Pharm Sci 2013;102(9):3037–57.
[15] Kass GE, Price SC. Role of mitochondria in drug-induced cholestatic injury. Clin Liver Dis 2008;12(1):27–51, [vii].
[16] Kaniwa N, Saito Y. Pharmacogenomics of severe cutaneous adverse reactions and drug-induced liver injury. J Hum Genet 2013;58(6):317–26.
[17] Corsini A, Ganey P, Ju C, Kaplowitz N, Pessayre D, Roth R, et al. Current challenges and controversies in drug-induced liver injury. Drug Saf 2012;35(12):1099–117.
[18] Wallis RM. Integrated risk assessment and predictive value to humans of non-clinical repolarization assays. Br J Pharmacol 2010;159(1):115–21.
[19] Gintant G. An evaluation of hERG current assay performance: translating preclinical safety studies to clinical QT prolongation. Pharmacol Ther 2011;129(2):109–19.
[20] Urban TJ, Goldstein DB, Watkins PB. Genetic basis of susceptibility to drug-induced liver injury: what have we learned and where do we go from here? Pharmacogenomics 2012;13(7):735–8.
[21] Chung WH, Hung SI, Hong HS, Hsih MS, Yang LC, Ho HC, et al. Medical genetics: a marker for Stevens-Johnson syndrome. Nature 2004;428(6982):486.
[22] Chen P, Lin JJ, Lu CS, Ong CT, Hsieh PF, Yang CC, et al. Carbamazepine-induced toxic effects and HLA-B*1502 screening in Taiwan. N Engl J Med 2011;364(12):1126–33.
[23] Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe'er I, Floratos A, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. Nat Genet 2009;41(7):816–9.
[24] Fontana RJ, Watkins PB, Bonkovsky HL, Chalasani N, Davern T, Serrano J, et al. Drug-Induced Liver Injury Network (DILIN) prospective study: rationale, design and conduct. Drug Saf 2009;32(1):55–68.
[25] Molokhia M, McKeigue P. EUDRAGENE: European collaboration to establish a case-control DNA collection for studying the genetic basis of adverse drug reactions. Pharmacogenomics 2006;7(4):633–8.
[26] Andrade RJ, Lucena MI, Fernandez MC, Pelaez G, Pachkoria K, Garcia-Ruiz E, et al. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. Gastroenterology 2005;129(2):512–21.
[27] Lucena MI, Molokhia M, Shen Y, Urban TJ, Aithal GP, Andrade RJ, et al. Susceptibility to amoxicillin-clavulanate-induced liver injury is influenced by multiple HLA class I and II alleles. Gastroenterology 2011;141(1):338–47.
[28] Benichou C, Danan G, Flahault A. Causality assessment of adverse reactions to drugs—II. An original model for validation of drug causality assessment methods: case reports with positive rechallenge. J Clin Epidemiol 1993;46(11):1331–6.

[29] Danan G, Benichou C. Causality assessment of adverse reactions to drugs–I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. J Clin Epidemiol 1993;46(11):1323–30.

[30] Rochon J, Protiva P, Seeff LB, Fontana RJ, Liangpunsakul S, Watkins PB, et al. Reliability of the Roussel Uclaf causality assessment method for assessing causality in drug-induced liver injury. Hepatology 2008;48(4):1175–83.

[31] Rockey DC, Seeff LB, Rochon J, Freston J, Chalasani N, Bonacini M, et al. Causality assessment in drug-induced liver injury using a structured expert opinion process: comparison to the Roussel-Uclaf causality assessment method. Hepatology 2010;51(6):2117–26.

[32] Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, et al. Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89(6):806–15.

[33] Watkins PB, Seeff LB. Drug-induced liver injury: summary of a single topic clinical research conference. Hepatology 2006;43(3):618–31.

[34] Daly AK, Day CP. Genetic association studies in drug-induced liver injury. Semin Liver Dis 2009;29(4):400–11.

[35] Cai Y, Yi J, Zhou C, Shen X. Pharmacogenetic study of drug-metabolising enzyme polymorphisms on the risk of anti-tuberculosis drug-induced liver injury: a meta-analysis. PLoS One 2012;7(10):e47769.

[36] Huang YS, Chern HD, Su WJ, Wu JC, Chang SC, Chiang CH, et al. Cytochrome p450 2E1 genotype and the susceptibility to antituberculosis drug-induced hepatitis. Hepatology 2003;37(4):924–30.

[37] Lee SW, Chung LSC, Huang HH, Chuang TY, Liou YH, Wu LSH. NAT2 and CYP2E1 polymorphisms and susceptibility to first-line anti-tuberculosis drug-induced hepatitis. Int J Tuberc Lung Dis 2010;14(5):622–6.

[38] Kim SH, Kim SH, Bahn JW, Kim YK, Chang YS, Shin ES, et al. Genetic polymorphisms of drug-metabolizing enzymes and anti-TB drug-induced hepatitis. Pharmacogenomics 2009;10(11):1767–79.

[39] Vuilleumier N, Rossier MF, Chiappe A, Degoumois F, Dayer P, Mermillod B, et al. CYP2E1 genotype and isoniazid-induced hepatotoxicity in patients treated for latent tuberculosis. Eur J Clin Pharmacol 2006;62(6):423–9.

[40] Aithal GP, Day CP, Leathart JB, Daly AK. Relationship of polymorphism in CYP2C9 to genetic susceptibility to diclofenac-induced hepatitis. Pharmacogenetics 2000;10(6):511–8.

[41] Pachkoria K, Lucena MI, Ruiz-Cabello F, Crespo E, Cabello MR, Andrade RJ, et al. Genetic polymorphisms of CYP2C9 and CYP2C19 are not related to drug-induced idiosyncratic liver injury (DILI). Br J Pharmacol 2007;150(6):808–15.

[42] Acuna G, Foernzler D, Leong D, Rabbia M, Smit R, Dorflinger E, et al. Pharmacogenetic analysis of adverse drug effect reveals genetic variant for susceptibility to liver toxicity. Pharmacogenomics J 2002;2(5):327–34.

[43] Daly AK, Aithal GP, Leathart JB, Swainsbury RA, Dang TS, Day CP. Genetic susceptibility to diclofenac-induced hepatotoxicity: contribution of UGT2B7, CYP2C8, and ABCC2 genotypes. Gastroenterology 2007;132(1):272–81.

[44] Lucena MI, Andrade RJ, Martinez C, Ulzurrun E, Garcia-Martin E, Borraz Y, et al. Glutathione S-transferase m1 and t1 null genotypes increase susceptibility to idiosyncratic drug-induced liver injury. Hepatology 2008;48(2):588–96.

[45] Watanabe I, Tomita A, Shimizu M, Sugawara M, Yasumo H, Koishi R, et al. A study to survey susceptible genetic factors responsible for troglitazone-associated hepatotoxicity in Japanese patients with type 2 diabetes mellitus. Clin Pharmacol Ther 2003;73(5):435–55.

[46] Choi JH, Ahn BM, Yi J, Lee JH, Nam SW, Chon CY, et al. MRP2 haplotypes confer differential susceptibility to toxic liver injury. Pharmacogenet Genomics 2007;17(6):403–15.

[47] Lang C, Meier Y, Stieger B, Beuers U, Lang T, Kerb R, et al. Mutations and polymorphisms in the bile salt export pump and the multidrug resistance protein 3 associated with drug-induced liver injury. Pharmacogenet Genomics 2007;17(1):47–60.

[48] Bhatnagar P, Day CP, Aithal G, Pirmohamed M, Bernal W, Daly AK. Genetic variants of hepatic transporters and susceptibility to drug induced liver injury. Toxicology 2008;253(1–3):10.

[49] Torgerson TR, Linane A, Moes N, Anover S, Mateo V, Rieux-Laucat F, et al. Severe food allergy as a variant of IPEX syndrome caused by a deletion in a noncoding region of the FOXP3 gene. Gastroenterology 2007;132(5):1705–17.

[50] Spraggs CF, Budde LR, Briley LP, Bing N, Cox CJ, King KS, et al. HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. J Clin Oncol 2011;29(6):667–73.

[51] Singer JB, Lewitzky S, Leroy E, Yang F, Zhao X, Klickstein L, et al. A genome-wide study identifies HLA alleles associated with lumiracoxib-related liver injury. Nat Genet 2010;42(8):711–4.

[52] Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenomics J 2008;8(3):186–95.

[53] Monshi MM, Faulkner L, Gibson A, Jenkins RE, Farrell J, Earnshaw CJ, et al. Human leukocyte antigen (HLA)-B*57:01-restricted activation of drug-specific T cells provides the immunological basis for flucloxacillin-induced liver injury. Hepatology 2013;57(2):727–39.

[54] Illing PT, Vivian JP, Dudek NL, Kostenko L, Chen Z, Bharadwaj M, et al. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. Nature 2012;486(7404):554–8.

[55] Alfirevic A, Pirmohamed M. Predictive genetic testing for drug-induced liver injury: considerations of clinical utility. Clin Pharmacol Ther 2012;92(3):376–80.

[56] Huyghe JR, Jackson AU, Fogarty MP, Buchkovich ML, Stancakova A, Stringham HM, et al. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nat Genet 2013;45(2):197–201.

[57] Grove ML, Yu B, Cochran BJ, Haritunians T, Bis JC, Taylor KD, et al. Best practices and joint calling of the HumanExome BeadChip: the CHARGE consortium. PLoS One 2013;8(7):e68095.

[58] Rafaels N, Huang LL, Leung DYM, Beck LA, Vergara CI, Boguniewicz M, et al. Exome chip genotyping reveals association with LAMA3 (laminin, alpha 3) gene and eczema herpeticum in a population of European descent. J Allergy Clin Immun 2013;131(2):Ab327–Ab.

[59] Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 2011;12(11):745–55.

[60] Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet 2011;12(8):529–41.

[61] Murphy SK, Yang H, Moylan CA, Pang H, Dellinger A, Abdelmalek MF, et al. Relationship between the methylome and transcriptome in patients with non-alcoholic fatty liver disease. Gastroenterology 2013;145(5):1076–87.

[62] Avery NJ, King RM, Knight S, Hourihane JO. Assessment of quality of life in children with peanut allergy. Pediatr Allergy Immunol 2003;14(5):378–82.

[63] Hautekeete ML, Horsmans Y, Van Waeyenberge C, Demanet C, Henrion J, Verbist L, et al. HLA association of amoxicillin-clavulanate–induced hepatitis. Gastroenterology 1999;117(5):1181–6.

[64] Stewart JD, Horvath R, Baruffini E, Ferrero I, Bulst S, Watkins PB, et al. Polymerase gamma gene POLG determines the risk of sodium valproate-induced liver toxicity. Hepatology 2010;52(5):1791–6.

[65] Lucena MI, Garcia-Martin E, Andrade RJ, Martinez C, Stephens C, Ruiz JD, et al. Mitochondrial superoxide dismutase and glutathione peroxidase in idiosyncratic drug-induced liver injury. Hepatology 2010;52(1):303–12.

This page is intentionally left blank

Chapter 23

Hemostasis and Anticoagulants

Matthew R. Wilson and R. Campbell Tait
Department of Haematology, Glasgow Royal Infirmary, UK

Chapter Outline

23.1 Overview 479
23.2 Hemostasis 479
  23.2.1 The "Cascade" Model of Hemostasis 480
  23.2.2 A Contemporary View of Hemostasis 481
  23.2.3 Regulation of Coagulation 481
  23.2.4 Fibrinolysis 482
23.3 Indications for Anticoagulation 482
  23.3.1 Venous Thromboembolism 482
  23.3.2 Stroke Prevention in Atrial Fibrillation 483
23.4 Parenteral Anticoagulants 484
  23.4.1 Unfractionated Heparin 484
  23.4.2 Low Molecular Weight Heparin 486
  23.4.3 Fondaparinux 487
  23.4.4 Danaparoid 487
  23.4.5 Direct Thrombin Inhibitors 487
23.5 Oral Anticoagulants 488
  23.5.1 Vitamin K Antagonists 489
  23.5.2 Direct Thrombin Inhibitors—Dabigatran Etexilate 491
  23.5.3 Factor Xa Inhibitors—Rivaroxaban and Apixaban 493
23.6 Conclusion and Directions for Future Research 494
References 494

23.1 OVERVIEW

The traditional anticoagulants in clinical use for the prevention or treatment of thromboembolic disease are heparin (and its analogs) and warfarin. However, they have major limitations: a narrow therapeutic index, high interindividual variability in a variable dose-response relationship that requires monitoring by laboratory testing, and numerous drug–drug interactions. These limitations have provided not only the impetus for pharmacogenomic studies of efficacy and toxicity for warfarin (described in Chapter 24), but also the development of other antithrombotic agents that target discrete factors in the coagulation cascade. In this chapter, we outline the coagulation cascade and describe the pharmacology of traditional and novel anticoagulants.

23.2 HEMOSTASIS

Blood coagulation and platelet-mediated primary hemostasis are essential physiological defense mechanisms against bleeding, in this way maintaining the integrity of the vascular system. Meanwhile, inherent controlling mechanisms prevent excessive coagulation and thus maintain vascular patency. The biochemistry behind blood coagulation has been carefully studied over the years, and much is known about the sequence of events that results in the formation of fibrin and therefore coagulation.

In normal circumstances, blood flows through the vasculature in a liquid phase until vascular injury occurs, and hemostasis must be initiated to temporarily seal the defect [1]. In the process of resolution of the injury, remodeling occurs with lysis of any excess thrombus. The systems involved are in a delicate physiological balance that, in health, is tipped in favor of an anticoagulant state whereby coagulation occurs only after appropriate stimulation. Factors that favor anticoagulation, thus maintaining the fluidity of blood, include [2]

- The constant flow of blood itself diluting any activated or prothrombotic factors
- A thin layer of plasma preventing platelets from direct exposure to the endothelium
- Expression by healthy endothelial cells of antiplatelet and anticoagulant factors in excess of procoagulant factors
- The requirement for multiple stimuli of platelets for full activation

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00023-2
© 2014 Elsevier Inc. All rights reserved.

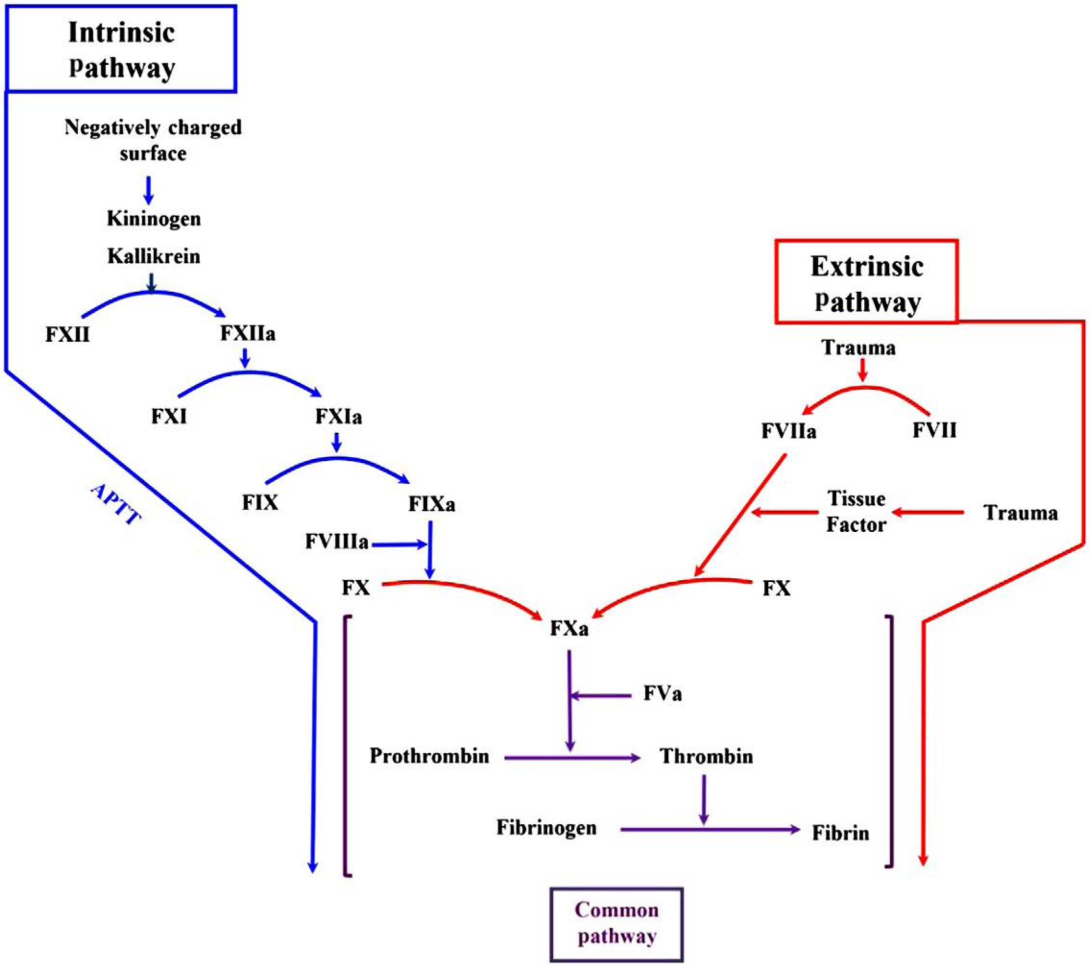


FIGURE 23.1 **Traditional "cascade" model of coagulation.** The extrinsic pathway (red) is initiated following damage to the blood vessel, when tissue factor is exposed and forms an activated complex with factor VIIa, which converts factor X to factor Xa. The intrinsic pathway (blue) is initiated when a negatively charged "foreign" surface is exposed to blood. Factor XII, in a complex with kininogen and kallikriein, becomes activated to form factor XIIa. Factor XIIa converts factor XI to factor XIa, which in turn activates factor IX. Factor IXa forms the "tenase" complex with factor VIIIa, which activates factor X to factor Xa. Both pathways are reflected in laboratory measurements—the prothrombin time (PT) simulates activity of the extrinsic pathway, and the activated partial thromboplastin time (APTT) simulates intrinsic activation. Both pathways converge at factor Xa, which marks the beginning of the common pathway (purple). Factor Xa, catalyzed by factor Va, converts prothrombin to thrombin. Thrombin causes production of fibrin from fibrinogen, which forms the basis of the blood clot. Note: Coagulation factors are denoted by "F" followed by the corresponding roman numeral.

- Circulating coagulation factors that exist as inactive zymogens or are bound to other cells and require activation by proteolysis or other means

23.2.1 The "Cascade" Model of Hemostasis

A comprehensive description of the steps involved in coagulation was first proposed in 1964 [3]. The "MacFarlane cascade"—also known as the waterfall model [4]—defines a stepwise series of enzymatic reactions whereby inactive zymogens are converted to biologically active enzymes (serine proteases). These activated coagulation factors activate each other in a cascade fashion culminating in the conversion of fibrinogen to fibrin, which polymerizes to form a fibrin clot.

According to MacFarlane, there are two distinct pathways that involve different activation stimuli and coagulation factors but eventually converge to a final "common" pathway (see Figure 23.1). Activation of the "intrinsic" pathway (so named because all of the components necessary are present in blood) is dependent on contact activation by a negatively charged surface and involves coagulation factors VIII, IX, XI, and XII. Its activity can be simulated and assessed *in vitro* by the activated partial thromboplastin time (APTT). The "extrinsic" pathway is activated when tissue factor is exposed to the circulation (for example in vascular trauma), and it involves tissue factor and factor VII. Similarly, it can be assessed *in vitro,* using the prothrombin time (PT). Both pathways result in activation of factor X, marking the beginning of the common pathway. Factor

Xa, catalyzed by Factor Va, then converts prothrombin to thrombin, which in turn enzymatically converts fibrinogen to fibrin. Interestingly, the role of platelets is considered unimportant in this model.

23.2.2 A Contemporary View of Hemostasis

In recent years, the "cascade" theory of blood coagulation has been disputed, and a modified cell-based model has been described that is thought to more accurately reflect the real-time events of hemostasis *in vivo* [5]. This new perspective builds on the classic cascade in the following ways [6,7]:

I. The cascade theory implies that clot formation can occur as a result of two independent pathways. It is now known that the interaction between factor VII (FVII) and tissue factor (TF) is the initiator of all coagulation processes (both natural and pathological).

II. The complex formed by TF and factor VII also participates in the activation of factor IX (as well as factor X), implying an element of cross-activation of the intrinsic pathway by the extrinsic system.

III. There is increasing recognition of the role of cells in coagulation. Different cells express varying levels of pro- and anticoagulant proteins and have different complements of receptors for components of hemostasis. It stands to reason, therefore, that cells have an important contribution to make in directing hemostasis *in vivo* [8]. In particular, activated platelets provide the phospholipid surface on which the coagulation factors can assemble—localizing the active coagulation pathway to the site of tissue injury.

IV. There are three consecutive phases to coagulation—initiation, amplification, and propagation. Platelets and thrombin are actively involved in the last two phases.

23.2.2.1 Initiation

It is now well recognized that exposure of blood to cells expressing TF results in the initiation of blood coagulation *in vivo* [1]. TF is a cell-bound transmembrane glycoprotein and a member of the class II cytokine family. It is constitutively expressed at biological boundaries such as skin, organ surfaces, vascular adventitia, and epithelial-mesenchymal surfaces, where it functions as a "hemostatic envelope." This ensures that, following disruption of vascular integrity, blood is immediately exposed to TF and coagulation is initiated. Endothelial cells themselves can express TF—in the healthy state, this is suppressed but upregulated when the endothelial cell is damaged.

Tissue factor forms a catalytic complex with factor VIIa on the phospholipid surface of the cell membrane, which then activates the zymogens factor IX and factor X. The activated factor X (factor Xa) then goes on to generate small amounts of thrombin.

23.2.2.2 Amplification

The levels of thrombin generated from the initiation phase are usually insufficient to generate significant fibrin formation. Thrombin's main function, however, is to activate factors V, VIII, and XI. In the amplification phase, factor XIa activates factor IX, which forms a complex with factor VIIIa. This complex ("tenase") activates sufficient factor Xa to form a complex with factor Va, termed the prothrombinase complex, production of which leads to the explosive generation of thrombin that ultimately leads to generation of a fibrin clot [1].

23.2.2.3 Propagation

The propagation phase relies on recruitment of activated platelets at the site of injury. Platelets provide a phospholipid surface that allows localization of the activated coagulation components for optimum generation of thrombin [2]. These components include the intrinsic tenase complex, the prothrombinase complex, and calcium. The "thrombin burst" brought about by these complexes leads to the generation of fibrin from fibrinogen to produce a stable fibrin clot. Thrombin also activates Factor XIII, which strengthens fibrin cross-linking and thrombin-activatable fibrinolysis inhibitor (TAFI), which protects the clot from plasmin-mediated fibrinolysis. (See Figure 23.2.)

23.2.3 Regulation of Coagulation

Each phase of the hemostatic pathway has associated inhibitory components. Once the TF-Factor VIIa complex stimulates the amplification process, it is rapidly inactivated by tissue factor pathway inhibitor (TFPI), which is released from thrombin-activated platelets found on the surface of endothelial cells [8]. Administration of intravenous heparin or subcutaneous low-molecular-weight heparin (LMWH) results in release of endothelial TFPI.

The most important inhibitor of coagulation is antithrombin, a serine protease inhibitor. It impedes fibrin generation by forming inactive complexes with thrombin and factor Xa (and to a lesser extent factors IXa and XIa). Heparin also greatly enhances the action of antithrombin.

Thrombin also proteolytically activates the anticoagulant protein C pathway. Thrombomodulin, a receptor expressed on endothelial cell surfaces, binds thrombin, resulting in the activation of protein C. Activated protein C, along with protein S as a cofactor, inactivates factors Va and VIIIa, therefore downregulating the amplification phase of coagulation.

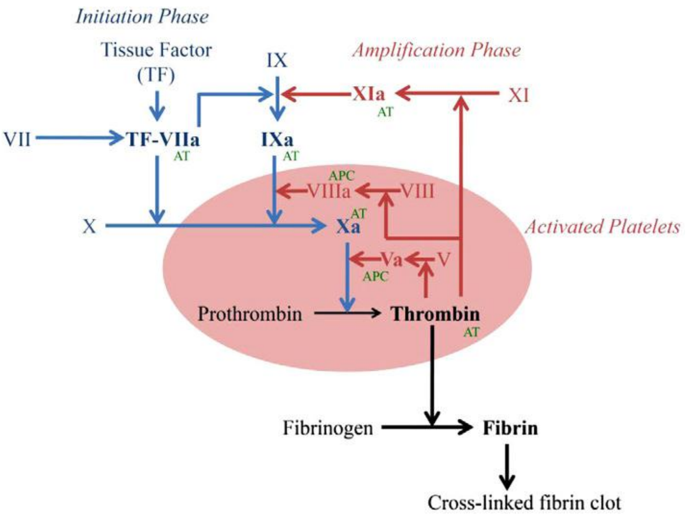


FIGURE 23.2 **Cell-based hemostasis model.** Exposure of blood to tissue factor is the initiating process of coagulation, with tissue factor forming a complex with factor VIIa. This complex activates factors IX and X, and factor Xa then generates small amounts of thrombin. Thrombin initiates a positive feedback network, activating factors V, VIII, and XI. Factors VIIIa and IXa form the "tenase complex," which activates factor X. Factor Xa forms the "prothrombinase complex" with factor Va which causes explosive generation of thrombin. Thrombin converts fibrinogen to fibrin, which forms the basis of the cross-linked blood clot. Activated platelets provide a surface for localization of the tenase-prothrombinase complex and are therefore essential for the "thrombin burst." Coagulation is moderated primarily by AT and APC. AT and APC are denoted in green and are positioned in the figure at the various points where they exhibit their inhibitory effects. Note: TF—tissue factor; AT—antithrombin; APC—activated protein C. Clotting factors are indicated by their respective roman numerals.

23.2.4 Fibrinolysis

Fibrinolysis describes the process of removing (lyzing) the clot formed by activation of hemostatic pathways, either in physiological response to vascular trauma or in pathological thrombosis. Plasmin is the principle mediator of fibrinolysis.

Fibrin itself can bind tissue plasminogen activator (tPA), greatly enhancing its ability to convert plasminogen to plasmin, the active enzyme. Plasmin is formed by proteolytic cleavage of plasminogen by both tPA and urokinase-type plasminogen activator (uPA). A serine protease itself, plasmin acts to dissolve the fibrin clot by splitting the cross-linked fibrin into a variety of small fibrin degradation products.

Regulation of the fibrinolytic system occurs via inhibition of tPA and plasmin by plasminogen activator inhibitor (PAI) and α2-antiplasmin, respectively.

23.3 INDICATIONS FOR ANTICOAGULATION

Anticoagulation, no matter how it is achieved, must only be initiated after careful consideration of its risks and benefits. The specific contraindications for individual anticoagulant agents are covered later in the chapter.

There are many clinical conditions in which anticoagulation has been shown to be beneficial; these are outlined in Table 23.1 [9]. In general terms, however, it is used for either primary or secondary prevention of venous or arterial thrombotic events. Venous thromboembolism and prevention of stroke in atrial fibrillation are the most common indications for anticoagulation. These are discussed in more detail in the following sections.

TABLE 23.1 Clinical Indications for Anticoagulant Treatment

Primary and secondary prevention of venous thromboembolism
Prevention of cardioembolic stroke in the presence of atrial fibrillation (AF) or cardiomyopathy
Cardioversion for arrhythmias such as atrial fibrillation
Valvular heart disease (patients suffering mitral stenosis or regurgitation who have AF or a history of systemic embolism, left atrial thrombus, or an enlarged left atrium)
Mechanical prosthetic heart valves
Bioprosthetic heart valves (conditions apply)
Acute coronary syndromes and ST-elevation myocardial infarction in the acute phase

23.3.1 Venous Thromboembolism

The following subsections discuss the epidemology, pathophysiology, and management of venous thromboembolism.

23.3.1.1 Epidemiology

Venous thromboembolism (VTE), which comprises deep vein thrombosis (DVT) and pulmonary embolism (PE), has a clinical incidence of approximately 1−1.5 per 1000 annually in adult populations [10]. Rates are slightly higher in men than in women. Two-thirds of cases present as DVT and one-third as PE with or without DVT. The major outcomes of venous thrombosis are death, recurrence, post-thrombotic syndrome, pulmonary hypertension, and major bleeding due to anticoagulation. Thrombosis is also associated with impaired quality of life, particularly when post-thrombotic syndrome develops [11]. It is estimated that within one month of an episode of DVT, 6% of patients will die; in the same period after an episode of PE, 10% will die. The mortality rate for PE has been estimated to be as high as 30% in autopsy-based studies, suggesting that PE is often not recognized before death.

VTE is rare before the age of 16, likely because the immature coagulation system is resistant to thrombosis. The risk of VTE then increases exponentially with age; this relationship is thought to relate to the increasing presence of other predisposing illnesses and also to an inherent increased coagulability with aging.

23.3.1.2 Pathophysiology

Virchow was the first to identify stasis of the blood, vessel wall injury, and hypercoagulability of the blood as the pathogenic triad responsible for thrombosis. It is thought that a combination of these factors is usually present. The major risk factors for VTE, and how they relate to Virchow's triad, are outlined in Table 23.2.

TABLE 23.2 **Risk Factors for VTE and Their Relationship to Virchow's Triad**

Virchow's Triad	Risk Factor
Stasis	Immobility
	"Travelers thrombosis"
	Previous VTE
	Surgery
	Trauma
	Obesity
Vessel wall injury	Surgery
	Trauma
Hypercoagulability	Age over 40
	Pregnancy
	Estrogen therapy
	Malignancy, including myeloproliferative neoplasm
	Genetic thrombophilia
	Antiphospholipid antibodies
	Obesity
	Surgery
	Trauma

Hypercoagulability can be either inherited or acquired. The most important inherited disorders associated with VTE result from the following:

- Defects in naturally occurring coagulation inhibitors—antithrombin, protein C, or protein S
- Resistance to activated protein C, usually due to a genetic mutation known as factor V Leiden (*F5* R506Q)

Deficiencies of antithrombin, protein C, and protein S are rare (combined prevalence of <1% in the general population), and so estimates of associated VTE risks are uncertain but have been quoted as creating a 10–40-fold increase in risk [12]. The factor V Leiden mutation is common, with an incidence of 5% in Caucasians, and carries a 5–7-fold increased risk of VTE [13].

23.3.1.3 Management

The mainstay of treatment for VTE is anticoagulation. Traditionally this has been achieved by initially administering low-molecular-weight heparin LMWH, followed by a course of oral anticoagulation. Urgent thrombolytic therapy may be used when the patient has hemodynamic compromise as the result of massive PE.

The optimum duration of anticoagulation in VTE is a contentious issue. Its purpose is to prevent recurrence; however, because this treatment carries a risk of bleeding, it should only continue if the risk of recurrent VTE, or its consequences, outweighs the risk of bleeding [14].

Standard practice has been to give three to six months of treatment for a first VTE, although this may be increased to "lifelong" if the patient is perceived to have a high risk of recurrence following an unprovoked VTE.

23.3.2 Stroke Prevention in Atrial Fibrillation

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia, increasing in prevalence with age [15]. AF accounts for approximately one-third of hospitalizations for cardiac rhythm disturbances. It has an estimated prevalence of 1.5–2%, but this is estimated to at least double in the next 50 years as the population ages [16].

The rate of ischemic stroke among patients with nonvalvular AF averages 5% per year, which is 2–7 times that for people without AF. This increased risk of stroke and systemic embolism is explained by a combination of pathophysiological mechanisms. The arrhythmia results in "flow abnormalities," as evidenced by blood stasis in the left atrium and in particular in the left atrial appendage. "Endocardial abnormalities" also occur, with progressive atrial dilatation, endocardial denudation, and fibroelastic infiltration of the extracellular matrix. Activation of hemostatic processes and platelets has also been described in AF, as well as inflammation and growth factor abnormalities [17]. These factors, either alone or in combination, result in an increased likelihood of fibrin thrombus formation and therefore embolic stroke [16].

Stroke in AF is often severe and often causes either long-term disability or death, resulting in a sizeable burden on health resources. An individual's predisposition to stroke in AF is related to a number of well-documented clinical and echocardiographic risk factors: age, hypertension, prior stroke or transient ischemic attack (TIA), diabetes, and structural heart disease.

The identification of these risk factors has led to the development of various stroke risk assessment schemes. The most recent and comprehensive of these risk factor stratification algorithms is the CHA_2DS_2-VASc scoring system, which is outlined in Table 23.3.

According to the European Society of Cardiology, if a patient has a score of 2 or more, an oral anticoagulant (OAC) should be prescribed; if the score is 1, either an oral anticoagulant or aspirin can be prescribed (an oral anticoagulant is preferred); if the score is 0, either aspirin can be prescribed or no antithrombotic therapy at all.

It must be remembered that, in considering an OAC, the advantages and disadvantages should be discussed with the patient (along with his or her preferences), and both the risk of bleeding complications and ability to safely sustain chronic stable anticoagulation should be evaluated.

23.4 PARENTERAL ANTICOAGULANTS

In the following sections, the structure and mechanism of action, pharmacokinetics, and other factors relating to unfractionated and low-molecular-weight heparin are discussed.

23.4.1 Unfractionated Heparin

Heparin was first isolated nearly 100 years ago by a second-year medical student, Jay McLean, who was working with the physiologist William Henry Howell at John Hopkins Medical School. McLean had been investigating procoagulant preparations when he came across a fat soluble phosphatide anticoagulant in canine liver tissue (hence the name heparin—"hepar," or "ήπαρ," is Greek for "liver"). Heparin was used in the 1930s as an intravenous anticoagulant in some human trials, and its use became widespread after the Second World War [18]. This original form of heparin is known as unfractionated heparin (UFH), mainly to distinguish it from derivatives that were developed more recently and are described later in the chapter.

23.4.1.1 Structure and Mechanism of Action

Heparin is a highly sulphated glycosaminoglycan, procured from tissues rich in mast cells such as porcine intestine or bovine lung. It is a polymer of repeating disaccharide units, with molecules ranging in molecular weight 3000–30,000 kDa (mean 15,000). Only approximately one-third of heparin molecules possess the unique pentasaccharide sequence that allows binding to the plasma protein antithrombin (AT)—this process is responsible for most of heparin's anticoagulant effect. The remaining two-thirds possess minimal anticoagulant activity at therapeutic concentrations, but at concentrations higher than usually obtained clinically, both high- and low-affinity

TABLE 23.3 The CHA_2DS_2-VASc Risk Factor–Based Scoring System for Stroke Risk in AF

Risk Factor	Condition	Points
C	Congestive heart failure/left ventricular dysfunction	1
H	Hypertension	1
A	Age ≥75	2
D	Diabetes mellitus	1
S	Stroke/TIA/thromboembolism	2
V	Vascular disease	1
A	Age 65–74	1
Sc	Sex category (female)	1

heparin catalyzes the antithrombin effect of a second plasma protein, heparin cofactor II [19].

The aforementioned pentasaccharide sequence binds to positively charged residues on AT, producing a conformational change that converts AT from a slow to a rapid inhibitor of serine proteases [20]. The resultant heparin/AT complex inactivates factor IIa (thrombin) and factor Xa. Binding of heparin to antithrombin occurs at the unique pentasaccharide sequence described above. Binding of this sequence alone is sufficient to generate anti-FXa and anti-thrombin activity. However, longer heparin chains containing at least 18 saccharides, including the pentasaccharide, have the additional capacity to bind thrombin thus localizing it to the activated AT and further enhancing thrombin inhibition. Thrombin inactivation prevents fibrin formation and also inhibits thrombin-induced activation of platelets and factors V, VIII, and XI.

23.4.1.2 Pharmacokinetics

Heparin is not absorbed orally and therefore must be administered parenterally. The two preferred routes of administration of UFH are by continuous IV infusion or subcutaneous (SC) injection.

At typical therapeutic doses, the half-life of intravenous UFH is 45–60 min, demanding continuous IV infusion for a sustained anticoagulant effect. Given subcutaneously, it has a lower bioavailability, with activity starting at 2 h and lasting approximately 10 h. For continued anticoagulation, therefore, subcutaneous dosing must be at least twice daily. If an immediate anticoagulant effect is required and heparin is to be administered subcutaneously, either a higher initial subcutaneous dose, or alternatively an IV bolus in conjunction with the first SC dose, should be given.

In the circulation, heparin binds nonspecifically to a number of plasma and cellular proteins other than AT, resulting in decreased bioavailability and substantial interpatient variability in anticoagulant response. It is for this reason that UFH requires frequent laboratory monitoring to assess the level of anticoagulation, as measured by the activated partial thromboplastin time (APTT) [21]. Clearance of heparin occurs via a rapid saturable and a much slower first-order mechanism combined. The saturable phase is thought to be due to binding to endothelial cell receptors and macrophages, which internalize and depolymerize heparin molecules. The slower, nonsaturable mechanism of clearance is predominantly renal. The complex kinetics of heparin clearance render the anticoagulant response to heparin nonlinear at therapeutic doses, with both intensity and duration of effect rising disproportionately with increasing dose.

23.4.1.3 Monitoring

Because of substantial interpatient variability in anticoagulant response to heparin, it is standard practice to monitor treatment and adjust the heparin dose based on the results of coagulation tests—usually using the APTT.

The evidence for adjusting the dose of heparin to maintain a “therapeutic range” is weak and based on a post hoc subgroup analysis of a descriptive study [20]. This is in stark contrast to the evidence for maintaining an international normalized ratio (INR) within a therapeutic range in patients on warfarin, which has been derived from consistent results of randomized trials and case-control studies.

Nevertheless, current practice uses a therapeutic APTT range of 1.5–2.5 times control for patients on unfractionated heparin, as this was shown to be associated with a reduced risk of VTE in the aforementioned retrospective study [20]. With this in mind, heparin dose adjustment nomograms are frequently used that give dose adjustment advice based on current APTT values. It should also be remembered that the measured response to the APTT varies among laboratory reagents and instruments. Therefore, the therapeutic APTT range at a particular institution should be adapted to the responsiveness of the reagent and coagulometer used [21].

23.4.1.4 Adverse Effects

As well as hemorrhagic complications common to all anticoagulants, heparin has some specific side effects. The most significant are heparin-induced thrombocytopenia (HIT) and osteoporosis. Rarer complications of treatment include skin reactions that can progress to necrosis, hyperkalemia, and hypersensitivity reactions.

Heparin-Induced Thrombocytopenia

One-third of patients have a slight fall in their platelet count (<30% compared with baseline) within the first four days of starting heparin. This is not associated with bleeding or thrombotic complications and does not require cessation of heparin. If the platelet count falls more markedly (>50% compared with baseline), then the possibility of heparin-induced thrombocytopenia (HIT) should be considered.

HIT is a life-threatening disorder that occurs in 0.3–6.5% of those receiving UFH. It is caused by the development of IgG antibodies against platelet factor 4 (PF4) and heparin. Patients with HIT have a thrombotic risk more than 30 times that of control populations. This paradoxical increased risk of thrombosis is explained by the fact that antibodies to PF4 and heparin cause platelet activation as well as depletion.

Thrombotic complications may affect any vascular bed and frequently occur at sites of vascular injury. Venous thromboses predominate in medical and orthopedic patients, whereas arterial and venous thromboses occur at a similar frequency in patients who have undergone cardiac or vascular surgery. Rarer sites of thrombosis include adrenal veins or cerebral venous sinuses.

Platelet counts rarely drop to a level that causes bleeding, and typically recover within 4–14 days after heparin is discontinued [22]. Despite this, the risk of thrombosis remains high for days to weeks after discontinuation of heparin, even if the platelet count has normalized. Atypical manifestations of the condition include heparin-induced skin necrosis, venous gangrene of the limbs, and anaphylactic-type reactions after receipt of an intravenous bolus of heparin.

Diagnosis of HIT is often made clinically, but this can be challenging as there are often other reasons that patients have reduced platelet counts, such as bacterial infection, drugs other than heparin, or bone marrow disease. Another challenging situation is the patient who has recently undergone cardiac surgery. The prevalence of heparin-dependent antibodies is high and thrombocytopenia is common in these patients, who are often on other drugs that could cause thrombocytopenia. The most helpful clinical diagnostic tool is observing a recovery in the platelet count after withdrawing all forms of heparin (including line flushes). Diagnosis can be confirmed in the laboratory using serologic or functional assays or both.

Management of HIT involves immediate cessation of all types of heparin and initiation of an alternative anticoagulant such as danaparoid or argatroban at therapeutic doses. Warfarin is contraindicated until the platelet count has recovered, as there have been reports of warfarin-induced skin necrosis and venous gangrene of the limbs. Platelet transfusion is relatively contraindicated during a HIT episode.

Osteoporosis

Osteoporosis can occur when heparin binds to osteoblasts, causing the release of factors that activate osteoclasts. This effect is normally only seen in prolonged use of UFH—2% of patients receiving treatment-dose UFH for more than nine months developed vertebral fractures [13].

23.4.1.5 Pharmacogenomics

Most pharmacogenetic analyses of heparin have been related to heparin-induced thrombocytopenia (HIT). There is some evidence that a weak genetic predisposition to HIT is related to a His/Arg polymorphism at amino acid 131 of the human FcγIIa receptor. One large study of this relationship found a statistically significant ($p = 0.03$) increase in the frequency of the His$_{131}$ genotype in patients with HIT when compared to healthy individuals [23]. More recently, it has been proposed that a polymorphism of the *GPIIIa* gene may modulate the prothrombotic effects of HIT. However, data on this association is not consistent. Patients with all Fc-receptor genotypes can develop HIT, and the association of these hemostatic polymorphisms with thromboembolic complications is certainly weak.

23.4.1.6 Reversal

Intravenous protamine sulphate can rapidly reverse the anticoagulant effects of heparin. Protamine sulphate is a basic protein derived from fish sperm that binds to heparin to form a stable salt. One milligram of protamine sulphate neutralizes approximately 100 units of heparin. Because the half-life of heparin is 60−90 min when given as an intravenous infusion, the total amount administered over the previous 2−2.5 h should be taken into account when calculating the protamine sulphate dose. However, when heparin has been given subcutaneously (and therefore is cleared with a longer half-life) prolonged or repeated administration of protamine sulphate may be required.

23.4.2 Low Molecular Weight Heparin

The realization that the anticoagulant action of UFH was mostly delivered by short pentasaccharides has led to the use of more efficient fractionation procedures, yielding a number of low-molecular-weight heparin (LMWH) preparations. Like UFH, LMWHs are heterogeneous with respect to molecular size and anticoagulant activity. They have a mean molecular weight of 4000−5000 with a molecular weight distribution of 1000–5000.

LMWHs, formed by chemical or enzymatic depolarization of UFH, have the following characteristics [19]:

- The smaller LMWH fragments have reduced ability to bind and therefore catalyze inhibition of thrombin but retain the ability to catalyze inhibition of factor Xa.
- Reduced nonspecific binding to plasma proteins, resulting in an improvement in the predictability of their dose-response relationship.
- Reduced binding to macrophages and endothelial cells, with a corresponding increase in their plasma half-life.
- Lower incidence of HIT, which is thought to be due to reduced binding to platelets and PF4.
- Weaker association with reduction in bone density, possibly due to reduced binding to osteoblasts.

Examples of LMWHs in clinical use are dalteparin, enoxaparin, and tinzaparin, which are typically administered in fixed or weight-adjusted doses for thromboprophylaxis and in weight-adjusted doses for therapeutic anticoagulation. Like heparin, they produce their major anticoagulant effect by catalyzing AT-mediated inhibition of coagulation factors. Of LMWH chains, 50–75% are too short to efficiently catalyze thrombin inhibition, and therefore their dominant effect is on AT-mediated inactivation of factor Xa.

LMWHs are cleared predominantly by the renal route, and therefore their biological half-life is increased in patients with renal failure. This longer half-life (3–6 h), when compared to UFH, means that a once-daily subcutaneous injection is sufficient for a sustained anticoagulant effect. Intravenous administration of LMWH is rarely used. In the setting of severe renal insufficiency in which therapeutic anticoagulation is required, use of UFH avoids the problems associated with impaired clearance of LMWH preparations.

Coagulation monitoring of LMWH is generally not necessary, but some authorities suggest that it may be appropriate in obese patients and those with renal insufficiency. The APTT is poorly sensitive to LMWH; however, monitoring is achieved by analyzing plasma anti-Xa activity. There is no specific creatinine clearance (CrCl) at which the risk of LMWH accumulation becomes clinically significant, but a value of 30 mL/min has been suggested in the literature [20]. If LMWH is used at CrCl below this level, anti-Xa monitoring and/or dose reduction should be considered.

If reversal of LMWH is required, protamine sulphate is administered in a dose of 1 mg per 100 anti-Xa units of LMWH (for example, 1 mg enoxaparin contains approximately 100 anti-Xa units and so requires 1 mg of protamine sulphate). A second, smaller dose of protamine sulphate can be given if bleeding continues.

To our knowledge, there are no published studies demonstrating a pharmacogenetic effect in LMWHs. However, it has been suggested that interleukin-1 receptor antagonist polymorphisms may be a useful marker to identify patients who benefit from LMWH in non-ST segment elevation acute coronary syndrome (ACS). In particular, carriers of this polymorphism showed a better control of von Willebrand factor increase in a small cohort of ACS patients [24].

23.4.3 Fondaparinux

The discovery that the minimum heparin fragment required for high-affinity binding to AT consists of a pentasaccharide, and that such pentasaccharides can be isolated, paved the way for the development of a relatively specific factor Xa inhibitor called fondaparinux.

23.4.3.1 Structure and Mechanism of Action

Fondaparinux is a synthetic drug consisting of the AT-binding pentasaccharide sequence found in the heparin molecule. It has a molecular weight of 1728 and binds to AT, producing conformational change at the reactive site of AT, resulting in a 300-fold increase in its inhibitory activity on factor Xa. This inhibition of factor Xa is irreversible; however, the interaction between fondaparinux and antithrombin is not, allowing each single molecule of fondaparinux to bind to, and activate, several molecules of antithrombin.

Because there are no additional saccharides to provide the necessary bridging function for thrombin binding, fondaparinux has no activity against thrombin.

23.4.3.2 Pharmacokinetics

Fondaparinux is administered as a subcutaneous injection and is rapidly and completely absorbed. Its half-life after SC injection is longer than that of LMWH (17 and 4 h, respectively). A steady state is reached after the third or fourth once-daily dose, and fondaparinux is excreted unchanged in the urine. It produces a predictable anticoagulant response with minimal nonspecific binding to plasma proteins other than AT.

Based on its long half-life and consistent anticoagulant effect, fondaparinux is administered as a once daily injection and does not require coagulation monitoring. Clearance is exclusively renal; thus, it is contraindicated in patients with a CrCl <30 mL/min.

23.4.3.3 Dosing and Reversal

Fondaparinux has largely replaced LMWH as the anticoagulant of choice in acute coronary syndromes after a study in 2006 [25] showed that it gave similar reduction in ischemic events but caused significantly less major bleeding. For this purpose, it is administered at a dose of 2.5 mg once daily. The same dose is used for thromboprophylaxis in the setting of orthopedic surgery, where it has been shown to be superior to prophylactic doses of enoxaparin in preventing VTE following major hip and knee surgery.

If fondaparinux is to be used for treatment of VTE, it is given at a dose of 7.5 mg for patients with body weight 50–100 kg, 5 mg for patients weighing <50 kg, and 10 mg in those weighing >100 kg.

Protamine sulphate does not bind to fondaparinux, so it cannot be used for reversal in the event of major bleeding.

23.4.3.4 Adverse Effects

Fondaparinux does not bind to platelet factor 4, and therefore has no capacity to cause HIT. Therefore, it can be safely used as an alternative anticoagulant in patients who have developed HIT, although it is unlicensed for this purpose. There is little clinical data available on the effects of fondaparinux on bone metabolism, but *in vitro* studies have suggested that it may not have the same osteoporotic effect as other heparin preparations.

23.4.4 Danaparoid

Danaparoid is a heparinoid that consists of the glycosaminoglycans heparin sulphate, dermatan sulphate, and chondroitin sulphate. It has low cross-reactivity with antibodies induced by heparin and is therefore primarily used as the anticoagulant of choice in patients with HIT. Danaparoid has a long half-life and has no specific antidote, so it can be problematic in patients requiring urgent surgery or those with major bleeding.

23.4.5 Direct Thrombin Inhibitors

Because thrombin is the central effector of coagulation and amplifies its own production, it is a natural target for direct pharmacologic intervention. Direct thrombin inhibitors (DTIs) bind with thrombin to prevent an interaction between the enzyme and substrates. The currently approved parenteral DTIs are bivalirudin and argatroban. Hirudin is no longer available.

The advantages of DTIs include a targeted specificity for thrombin, the ability to inactivate clot-bound thrombin, and an absence of plasma protein and platelet interactions that are responsible for problems with heparin, such as HIT. DTIs exert their anticoagulant effect, independent of antithrombin, by binding directly to thrombin and forming an inactive stoichiometric complex. They do not bind plasma proteins and so have a more predictable anticoagulant effect than heparin [20].

There are no specific antidotes or reversing agents for DTIs; however, their short half-lives mean that such agents should be rarely required. Hemodialysis or hemoperfusion can remove bivalirudin or argatroban.

23.4.5.1 Bivalirudin

Bivalirudin has a smaller molecular weight than the other DTIs and is given intravenously. It has a shorter half-life (25 min), and its elimination is only partially renal. Bivalirudin's mechanism of action is similar to that of other DTIs, the only difference being that the bond it forms with thrombin is reversible. It is licensed as an alternative to heparin in patients undergoing percutaneous interventions for acute coronary syndromes and in patients with HIT.

23.4.5.2 Argatroban

Argatroban is a small-molecule competitive inhibitor of thrombin given intravenously, with a half-life of 40–50 min. It requires anticoagulation monitoring of the APTT. Argatroban is metabolized by the liver cytochrome P450 3A4/5 enzyme system and consequently must be used with caution in patients with hepatic dysfunction. Because it is not renally excreted, it is useful for patients with HIT and severe renal impairment. (See Figure 23.3.)

23.5 ORAL ANTICOAGULANTS

For many decades, the field of oral anticoagulation has been dominated by vitamin K antagonists, specifically warfarin. This section provides an overview of vitamin K antagonists as well as some of the newer anticoagulants that have progressed through clinical trials in the last few years and are now approved for clinical use—dabigatran, rivaroxaban, and apixaban.

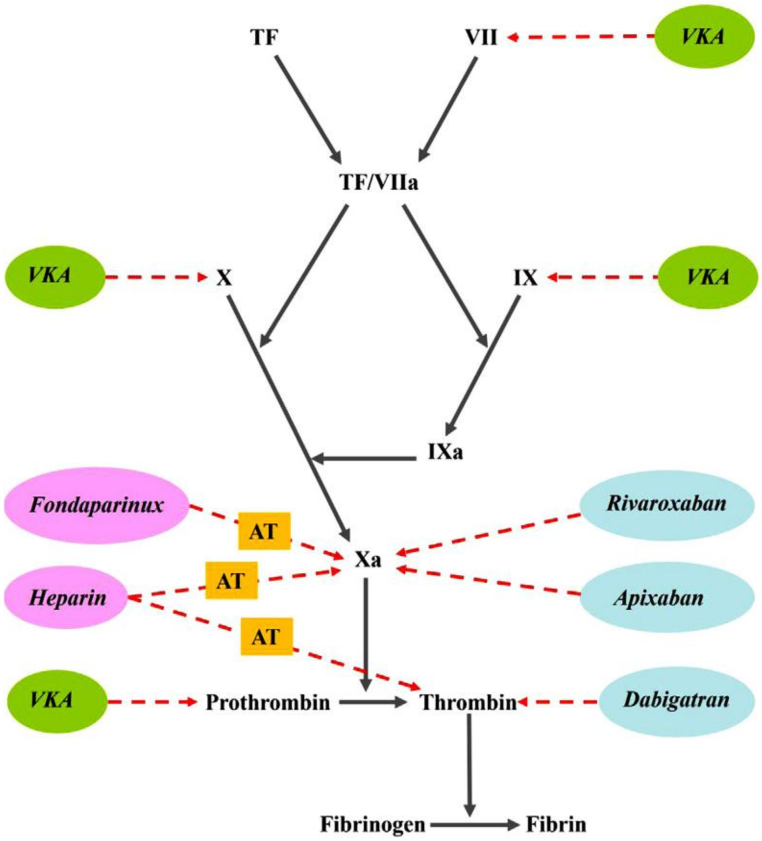


FIGURE 23.3 Anticoagulant drugs and their targets in the coagulation process. Red dotted lines denote inhibition of clotting factors by each anticoagulant. TF—tissue factor; VKA—Vitamin K Antagonist; AT—Antithrombin.

23.5.1 Vitamin K Antagonists

The vitamin K antagonists (VKAs) were first discovered by Karl Paul Link at the University of Wisconsin in the 1930s. Cattle in the Northern United States and Canada had been afflicted by an outbreak of an unusual disease characterized by fatal bleeding either spontaneously or from minor injuries. Investigators discovered that moldy silage made from sweet clover (*Melilotus alba* and *Melilotus officinalis*) was implicated, and that it contained a hemorrhagic factor that reduced the activity of prothrombin. Link, along with his student, Harold Campbell, then discovered that this anticoagulant in sweet clover was 3,3′-methylenebis (4-hydroxycoumarin). In 1948, work by Link led to the synthesis of warfarin, which was initially used as a rodenticide in 1952 but was then approved for human use in 1954. The name *warfarin* is derived from WARF (Wisconsin Alumni Research Foundation) and *arin* from coumarin [26]. The pharmacology of warfarin is briefly discussed here; pharmacogenomics are discussed in detail in Chapter 24.

The effectiveness of VKAs in primary and secondary prevention of venous thromboembolism and in prevention of systemic embolism in patients with prosthetic heart valves or atrial fibrillation has been well established by robust clinical trials. However, their widespread use also presents several challenges in clinical practice.

There are two broad classes of VKA—the coumarins and the indandiones. Warfarin is the most widely used coumarin, other forms being acenocoumarol and phenprocoumon, which have respectively shorter and longer half-lives than warfarin. Examples of indandiones are phenindione and fluindione. Because warfarin is by far the most commonly used VKA worldwide, it is used interchangeably with VKA in this discussion.

23.5.1.1 Mechanism of Action

Warfarin's mechanism of action is covered in more detail in Chapter 24. In summary, it prevents the carboxylation of clotting factors II, VII, IX, and X by inhibiting vitamin K epoxide reductase. These poorly carboxylated clotting factor forms have reduced ability to bind calcium and localize the coagulation process to phospholipid surfaces, and are therefore biologically inactive [13].

Paradoxically, VKAs also inhibit carboxylation of the regulatory anticoagulant proteins C, S, and Z and therefore can cause a transient procoagulant effect when initiated in the acute phase of a thrombotic event. Once the vitamin K–dependent clotting factor levels are reduced, the anticoagulant effect outweighs any procoagulant activity. It is for this reason that, in the treatment of acute VTE, it is advised that an additional anticoagulant (usually heparin) be given for at least five days (and until the INR is ≥2 for at least 24 h), by which time any procoagulant phase will have passed [9].

23.5.1.2 Pharmacokinetics

Warfarin is highly water soluble, is rapidly absorbed from the gastrointestinal tract, and has high bioavailability, reaching maximal blood concentrations within about ≈90 min of oral administration. It circulates bound to plasma proteins, mainly albumin (98–99%), and is metabolized by the hepatic cytochrome P450 enzymatic system [13].

Warfarin is actually a racemic mixture of two optically active isomers, known as the R and S enantiomers. These enantiomers have different anticoagulant potency and are metabolized in slightly different ways. The S enantiomer is 2.7–3.8 times more potent than the R enantiomer and undergoes approximately 90% oxidative metabolism, primarily by the CYP2C9 enzyme of the cytochrome P450 system and, to a lesser extent, by the CYP3A4 enzyme. The R enantiomer undergoes approximately 60% oxidative metabolism by the CYP1A2 and CYP3A4 cytochrome P450 enzymes [27].

23.5.1.3 Interactions

It is well documented that warfarin is susceptible to many interactions with drugs or food, and this constitutes its major disadvantage as a therapeutic agent. Genetic factors also play an important role in warfarin metabolism and therefore can determine an individual's response to the drug. This is discussed in a separate chapter.

The list of drugs that have been shown to interact with warfarin is long and ever expanding. Manufacturer product information for warfarin lists >200 specific agents that may interfere with it. A systematic review of available literature on the topic was carried out in 2005, and warfarin–drug interactions were rated according to interaction direction, clinical severity, and quality of evidence [28]. The investigators listed 72 individual drugs that they deemed as having either a "highly probable" or a "probable" clinically significant interaction with warfarin. Table 23.4 outlines some of the different ways in which warfarin–drug interactions can come about, as well as examples of offending drugs.

Warfarin is also susceptible to interactions with some foods and herbal remedies, which may be problematic for clinicians because patients are less likely to disclose intake of such products. Excessive use of alcohol is also known to affect the metabolism of warfarin and can elevate the INR. Many herbs and spices can interact with warfarin, such as ginger, garlic, ginseng, and ginkgo biloba. St John's Wort, sometimes used in depression, induces hepatic enzymes and therefore has an inhibitory effect on warfarin. In 2003, the UK Committee on Safety of Medicines received several reports of increased INR in patients taking warfarin and cranberry juice [29]. To date, a causal relationship has not actually been proven, but it is generally accepted that warning patients to avoid this combination is advisable.

TABLE 23.4 Mechanisms for Drug Interactions with Vitamin K Antagonists, with Examples

Mechanism	Example
Inhibition	
Reduced absorption due to binding	Cholestyramine
Induction of hepatic metabolism	Barbiturates, rifampicin, azathioprine, carbamazepine
Enhancement of vitamin K pathway	Vitamin K, lipid emulsions
Potentiation	
Inhibition of cytochrome P450 (CYP3A4)	Clarithromycin
Inhibition of cytochrome P450 (CYP2C9), S enantiomer	Sulfinpyrazone, metronidazole
Stereoselective inhibition of hydroxylation, R enantiomer	Cimetidine, omeprazole
Reduced nonstereoselective clearance (R and S enantiomers)	Amiodarone
Inhibition of cyclic interconversion of vitamin K	Second- and third-generation cephalosporin antibiotics
Increased metabolism of coagulation factors	Thyroxine, androgens
Potentiation of warfarin receptor effect	Clofibrate
Inhibition of platelet function	Aspirin, nonsteroidal anti-inflammatory drugs, high-dose penicillin
Increased absorption of warfarin	Acarbose

23.5.1.4 *Monitoring the Anticoagulant Effect of Warfarin*

Monitoring of warfarin's anticoagulant effect is perhaps more important than monitoring any other anticoagulant's effect because of its narrow therapeutic window and unpredictable pharmacokinetics and interactions. Prothrombin time (PT) is the most commonly used test for monitoring VKA therapy. The PT is particularly sensitive to reduced factor VII levels, but is also prolonged by reduced factors II and X. The result (in seconds) for a prothrombin time varies according to the type of analytical system used because of variations between different batches of thromboplastin used in the test reagent. A standardized calibration system called the International Normalized Ratio (INR) was adopted in 1982 [30]. The INR is the ratio of a patient's prothrombin time to a mean normal PT (derived from normal samples), raised to the power of the International Sensitivity Index (ISI) of the analytical system used.

$$\text{INR} = \left(\text{PT}_{\text{patient}}/\text{mean normal PT}\right)^{\text{ISI}}$$

The "target" INR for an individual patient depends on the reason for anticoagulation. For example, the first episodes of VTE should be treated with an INR target of 2.5, as should episodes of AF for prevention of stroke. There are some situations where higher target ranges are used, such as recurrent VTE despite anticoagulation (target 3.5) or mechanical heart valves (target 3 or 3.5 depending on the type of valve used).

Initiation of warfarin is normally guided using dosing nomograms, with higher loading doses adjusted for patient age and subsequent doses adjusted according to the INR. If rapid anticoagulation is not required, a "slow-loading" regimen has been shown to be safe and normally achieves therapeutic anticoagulation at 3–4 weeks [9].

It is a general experience, confirmed by some studies [31], that the quality of anticoagulation control is higher and the rate of bleeding lower when patients are monitored by dedicated anticoagulation clinics. Such clinics have specialized and experienced staff who promote patient education and use computer programs to help optimize anticoagulant therapy.

23.5.1.5 *Management of Bleeding and Reversal*

Bleeding is the most important complication of anticoagulation therapy and, because of the widespread use of warfarin as an oral anticoagulant, it is a fairly common clinical scenario. The annual risk of bleeding on VKAs in prospective studies has been reported to be 0.1–1.0% for fatal episodes, 0.5–6.5% for major episodes, and 6.2–21.8% for minor bleeding. The risk of bleeding increases exponentially as the achieved INR increases beyond 4.5.

There are several strategies for the management of bleeding or overanticoagulation in patients on a VKA. Which is used depends on the degree of overanticoagulation and the severity of the bleeding. Guidelines from the British Committee for Standards in Haematology on this matter are summarized in Table 23.5 [9].

Vitamin K (phytomenadione) works faster when administered intravenously, causing initial reductions in INR at 2 h. Correction of the INR to the normal range is generally achieved at 8–24 h if hepatic function is normal and a sufficient dose is given. Anaphylactoid reactions may occur with IV phytomenadione, especially if large IV doses are administered rapidly. The frequency of this complication is thought to be about 3 per 10,000 doses administered [27].

TABLE 23.5 Management Strategies in the Overanticoagulated or Bleeding Patient on a VKA

Clinical Scenario	Management
INR >5.0 with no bleeding	Withold 1–2 doses of VKA and reduce maintenance dose
INR >8.0 with no bleeding	1–5 mg of oral vitamin K
Major bleeding	25–50 units/kg four-factor prothrombin complex concentrate and 5 mg intravenous vitamin K
Nonmajor bleeding	1–3 mg intravenous vitamin K
Emergency surgery (that can be delayed by 6–12 h)	2–5 mg intravenous vitamin K
Emergency surgery (that cannot be delayed)	25–50 units/kg four-factor prothrombin complex concentrate and 5 mg intravenous vitamin K

Source: From Keeling et al. [9].

Prothrombin complex concentrates consist of the coagulation factors II, VII, IX, and X. They are extremely effective at rapid and complete reversal of anticoagulation in the warfarinized patient. They do not require a cross-match and can be infused over 15–30 min. These plasma products undergo several viral inactivation steps during their manufacture, thus minimizing the risk of viral transmission. The other main potential adverse effect is thrombosis, although this risk is also small.

23.5.1.6 *Nonhaemorrhagic Adverse Effects*

Rare side effects of VKAs are skin necrosis, purple toe syndrome, alopecia, rash, and toxic hepatitis.

23.5.2 Direct Thrombin Inhibitors—Dabigatran Etexilate

The clinical burden of thromboembolic disease and the limitations of current therapies have prompted the development of new anticoagulant agents that directly target a single enzyme in the coagulation cascade. A number of such agents have progressed through clinical development over the last few years, with those first to be licensed including dabigatran, rivaroxaban and apixaban.

Dabigatran is a selective direct thrombin inhibitor given as dabigatran etexilate (an orally absorbable prodrug), as dabigatran itself is not absorbed from the gut. It was first approved for use in VTE prevention after elective total knee or hip arthroplasty [32], but it has now been shown to be superior to warfarin in stroke prevention in nonvalvular AF [33] and noninferior to warfarin in treatment of VTE [34]. Along with the other new oral anticoagulant agents, it represents an alternative to warfarin for long-term anticoagulation with the benefit of fewer interactions and no requirement for monitoring.

23.5.2.1 *Pharmacology*

Dabigatran does not require a cofactor (unlike indirect anticoagulants such as heparin), and it directly inhibits both free and clot-bound thrombin. It is a small synthetic molecule (471.5 d) that prevents access to the active site on thrombin by forming a salt bridge via hydrophobic interactions [35]. Dabigatran also prevents thrombin-induced platelet aggregation.

23.5.2.2 *Pharmacokinetics*

Dabigatran has a predictable pharmacokinetic profile, allowing for a fixed-dose regimen that has been found to be most effective in clinical trials when given twice daily. Systemic bioavailability of oral dabigatran etexilate has been estimated at 6–7% in healthy elderly subjects [36]. The prodrug dabigatran etexilate is rapidly converted (after oral administration and hepatic processing) to dabigatran, with peak plasma dabigatran concentrations recorded approximately 1.5 h after ingestion [37]. After the peak has been reached, levels fall in a biphasic manner consistent with a rapid distribution phase, resulting in a more than 70% decrease 4–6 h after ingestion [36].

Although part of the bioconversion from prodrug to active metabolite occurs in the liver, the cytochrome P450 system is not involved; thus, dabigatran is not susceptible to interference from inducers or inhibitors of these liver enzymes. With repeated dosing, the terminal elimination half-life is 12–17 h and it takes 2–3 days to reach steady-state levels [27]. Of the administered dose, 85% is excreted by renal clearance, almost all as unchanged dabigatran, with the remaining 15% being conjugated to form pharmacologically active compounds. It is for this reason that dabigatran should be avoided in patients with severe renal impairment and that moderate renal impairment be accounted for when choosing the dosing regimen.

23.5.2.3 *Interactions*

Although dabigatran metabolism is not affected by changes in cytochrome P450 enzyme expression, it does display P-glycoprotein (P-gp) efflux transporter–mediated interactions. For example, coadministration with rifampicin (a P-gp inducer) decreases peak dabigatran concentrations and therefore drug efficacy [38]. The strong P-gp inhibitors quinidine and ketoconazole are contraindicated because they markedly increase exposure to dabigatran. There have been reports of interactions with other P-gp inhibitors such as amiodarone and verapamil, but the increase in bleeding risk is thought to be mild and may only be important if combined with old age or reduced creatinine clearance [27].

23.5.2.4 Monitoring the Anticoagulant Effect of Dabigatran

Dabigatran, rivaroxaban, and apixiban have been shown in randomized trials to be effective and safe without monitoring [32–34]. This lack of a need for monitoring is based on the assumed similarity in pharmacokinetic and pharmacodynamic responses between individuals within a relatively wide therapeutic window. It is worth remembering, however, that clinical trials often exclude patients with impaired renal function, children, the very elderly, those with bleeding risk factors, and those at extremes of body weight. It may be that more experience is needed in the use of these drugs in such patient groups before it can be safely assumed that monitoring is not required. There are still instances, however, when rapid measurement of the concentration of an anticoagulant or the degree of anticoagulation is necessary. Examples include hemorrhage, suspected overdose, emergency surgery and compliance assessment.

Traditional anticoagulant tests have limitations for measuring dabigatran effect—the PT is relatively insensitive, the relationship between the APTT and dabigatran concentration is nonlinear, and the usual form of TCT (thrombin clotting time) is oversensitive. It is suggested that each laboratory know the sensitivity of their own APTT reagent to dabigatran and advise on interpretation. Although TCT is oversensitive, a result in the normal range is helpful as it suggests that the level of dabigatran is likely to be very low [39].

Specific chromogenic anti IIa assays are available for the measurement of drugs that inhibit thrombin [40,41]. Ecarin clotting time (ECT) assays can also be used as a direct measure of direct thrombin inhibitors. This test has a linear dose-response relationship throughout the range of dabigatran concentrations expected during prophylaxis or therapeutic doses, but it is not always routinely available [27].

A genome-wide association study of 2944 participants of the Randomized Evaluation of Long-term Anticoagulation Therapy (RE-LY) study [45] identified a single-nucleotide polymorphism (SNP), rs2244613, in the *CES1* gene to be associated with trough concentrations, and the SNPs in *ABCB1* (rs4148738) and *CES1* (rs8192935) to be associated with peak concentrations at genome-wide significance. The carriage of the *CES1* rs2244613 minor allele occurred in 32.8% of patients in RE-LY and was associated with lower exposure to active dabigatran metabolite and with a lower risk of bleeding. This raises the possibility that routine genotyping may enable clinicians to tailor the dose of dabigatran for individual patients and thereby optimize the balance between efficacy and safety.

Bleeding and Reversal of the Anticoagulant Effect of Dabigatran

In the RE-LY trial, which compared dabigatran to warfarin in prevention of stroke for patients with AF, dabigatran showed similar incidences of major bleeding when compared to warfarin and caused fewer intracranial bleeds. It did, however, cause a dose-dependent increase in the incidence of major GI bleeding, which was statistically significantly greater than the risk with warfarin [33].

Perhaps the most significant limitation of the new oral anticoagulant agents is the lack of clinical evidence to guide management when rapid reversal of anticoagulation is required. As yet, dabigatran has no specific antidote, although products are in development. Its relatively short half-life means that, in situations where reversal within 24–48 h is acceptable (for example, prior to surgery), cessation of the drug and allowing it to be cleared may be appropriate. In patients with normal renal function, discontinuation of dabigatran 24 h before surgery decreases plasma levels to approximately 25% of steady-state trough levels, reducing to 5–10% 48 h after discontinuation [42].

In the event of bleeding complications in patients receiving dabigatran, there are a number of potential options for reversal of anticoagulation:

- *Recombinant factor VII*—this procoagulant and general haemostatic agent can initiate hemostasis by directly activating thrombin on the surfaces of platelets in the absence of tissue factor. *In vitro* and animal studies suggest that it can antagonize the effect of a number of anticoagulants, including dabigatran [42]. Clinical data and real-life experience is still lacking, however.
- *Prothrombin complex concentrates (PCCs)*—animal models have shown that PCCs can significantly reduce the effects of dabigatran on bleeding time [43]. A randomized controlled *in vivo* human study in healthy volunteers demonstrated that administration of a PCC resulted in virtually complete correction of coagulation after intake of rivaroxaban, but not after intake of dabigatran. It has been postulated that, because only one dosing regimen of PCC was used in this study, it may be that different doses may show efficacy for dabigatran reversal as well [44].
- *Hemodialysis and hemofiltration*—dabigatran is dialyzable because of its relatively low plasma-protein binding. Hemodialysis is therefore a potential option for rapid reversal of dabigatran's anticoagulant effect, especially in patients with renal impairment. In a study of 12 subjects with end-stage renal failure who received 50 mg of dabigatran, hemodialysis removed 62% of circulating dabigatran within 2 h [42].

Transfusion of fresh frozen plasma (FFP) is thought to have little benefit in dabigatran-associated bleeding. It is of benefit in situations where there is a depletion of clotting factors, but dabigatran causes inhibition rather than depletion of clotting factors, and FFP is unable to overcome this.

23.5.2.5 Nonhaemorrhagic Side Effects

Dyspepsia is the most frequently reported and clinically significant nonhemorrhagic side effect of dabigatran. Approximately 11% of patients experiencing dyspepsia with dabigatran have been reported in large clinical trials, compared to 6% with warfarin [33]. One of the earlier oral direct thrombin inhibitors, ximelegatran, was withdrawn because of severe liver injury. Subsequently, a genome-wide association study identified SNPs in *HLA-DRB1*0701* and *DQA1*0201* to be strongly associated with this, suggesting a possible immune pathogenesis [46]. To date, no dabigatran study has shown liver toxicity to be a clinically significant side effect. One large trial reported a small increased risk of myocardial infarction with dabigatran compared to warfarin (relative risk 1.35) [33], but the explanation for this is unclear.

23.5.3 Factor Xa Inhibitors—Rivaroxaban and Apixaban

Rivaroxaban and apixaban are direct inhibitors of factor Xa that have been evaluated in large, randomized control trials for the prevention and treatment of thromboembolic disease. Both drugs have been found to be either noninferior (rivaroxaban) or superior (apixaban) to warfarin in prevention of stroke in AF, with fewer intracranial bleeding events [47,48]. Similarly, they have been evaluated for treatment and prevention of venous thromboembolism and found to be noninferior to warfarin while causing less bleeding [49,50]. Like dabigatran, these drugs have more predictable pharmacokinetics and fewer interactions than warfarin, and do not require monitoring.

23.5.3.1 Mechanism of Action

Rivaroxaban and apixaban share the same mode of action. They are oral, highly selective, direct competitive inhibitors of factor Xa. Also, they are specific for factor Xa and do not inhibit related serine proteases, including thrombin, plasmin, factor VIIa, factor IXa, or activated protein C [51]. Both drugs reversibly bind to the active site of factor Xa, therefore suppressing the generation of new thrombin molecules. Their mechanism of action differs from indirect factor Xa inhibitors, such as fondaparinux, in that they inhibit both free and fibrin-bound factor Xa and do not rely on the presence of antithrombin [52].

23.5.3.2 Pharmacokinetics

Rivaroxaban and apixaban display similar and predictable pharmacokinetic and pharmacodynamic profiles. They are absorbed rapidly, with maximum plasma concentrations achieved at 4h after administration [53]. Apixaban can be taken without food, but higher doses of rivaroxaban (>15mg) should be taken with food to allow complete absorption. Rivaroxaban is dosed once daily except in treatment of acute VTE, when it is taken twice daily for the first three weeks before reverting to a once-daily regimen. Apixaban is administered twice daily.

Rivaroxaban is partially metabolized by the CYP450 system, specifically through the CYP3A4/5 and CYP2J2 pathways [51]. Apixaban is metabolized via CYP3A4-dependent mechanisms with minor contributions by CYP1A2 and 2J2. Both drugs are substrates of P-glycoprotein. Patients with hepatic dysfunction may have impaired metabolic activation, so rivaroxaban and apixaban should not be used in patients with moderate or severe hepatic disease (Child Pugh B/C).

Rivaroxaban has a half-life of 5–13h and has a dual mode of elimination: one-third is eliminated by the hepatobiliary route in feces; the other two-thirds are excreted in the urine. Rivaroxaban's half-life is prolonged by reduced renal function; therefore, caution is warranted in patients with moderate renal impairment and dose reduction is recommended for those with a CrCl of 15–49mL/min. It is contraindicated in patients with severe renal dysfunction (CrCl <15mL/min). Apixaban also displays a dual mode of elimination, but only 25% is excreted renally and therefore is less susceptible to accumulation in renal dysfunction. Nevertheless, it is still recommended that apixaban is avoided in patients with severe renal impairment [51].

23.5.3.3 Interactions

Both drugs are susceptible to drug interactions as they are metabolized through CYP450 and are P-gp substrates. Concomitant use of strong CYP3A4 inhibitors and P-gp inhibitors, such as the azole antifungals or protease inhibitors, is not recommended as it may lead to over anticoagulation [51]. Less potent inhibitors of these systems may be used without any dose adjustment, but any medication that is known to inhibit both pathways should be evaluated for use. Similarly, coadministration with strong inducers of CYP3A4 and P-gp should be avoided (examples include rifampicin, phenytoin, carbamazepine, St John's Wort).

As with all anticoagulants, vigilance with concurrent use of antiplatelet or other antithrombotics is warranted to avoid potentiation of the anticoagulant effect.

23.5.3.4 Monitoring of Anticoagulant Effect

Rivaroxaban and apixaban have been shown to be safe for use without routine monitoring because of their predictable pharmacokinetic characteristics. Both drugs cause only minor prolongation of the clotting times of APTT and PT assays [39]. Rivaroxaban has a more pronounced effect on APTT and PT than does apixaban, and PT is generally more sensitive to both drugs than is APTT. Again, it is advisable that each individual laboratory know the sensitivity of its PT and APTT assays to factor Xa inhibitors.

Chromogenic assays provide a more reliable measure of plasma concentrations of direct factor Xa inhibitors. The assays used have different dynamic ranges for rivaroxaban and apixaban, so a product-specific calibrator should be used.

23.5.3.5 *Bleeding and Reversal of Anticoagulant Effect*

Significant forms of bleeding have been found to be statistically lower with factor Xa inhibitors than with warfarin. The new oral anticoagulants as a group have shown significantly higher rates of gastrointestinal hemorrhage, but subgroup analyses have suggested that rates are lower in factor Xa inhibitors when compared to dabigatran [54].

At present there is no specific antidote available to antagonize the anticoagulant effects of either rivaroxaban or apixaban; however, such products are in development. In the case of overdose, activated charcoal may be used to reduce absorption.

Unlike dabigatran, renal dialysis is unlikely to be of benefit in enhancing the clearance of rivaroxaban or apixaban due to high plasma-protein binding. Possible alternative strategies include cessation of treatment and supportive management, or administration of prothrombin complex concentrates or recombinant factor VIIa in the case of severe bleeding. Limited data exist to support the use of these products other than some preclinical work in healthy volunteers [44] showing normalization of coagulation tests after their administration.

23.5.3.6 *Nonhemorrhagic Adverse Effects*

The most frequently reported adverse effects in clinical trials evaluating rivaroxaban and apixaban were nausea, vomiting, and constipation. The number of patients with elevated liver enzymes was low, and there were no associated increases in cardiovascular events.

23.6 CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

The field of anticoagulation has undergone significant change in recent years, with the advent of new oral anticoagulants believed by many to have the potential to replace warfarin entirely. Warfarin is very well characterized, having had over 50 years of widespread use in healthcare. However, its pharmacology continues to cause clinical problems, with many patients either not receiving optimum anticoagulation or suffering bleeding events as a result of overanticoagulation.

The new, target-specific anticoagulants have a rapid onset of action, low potential for food and drug interactions, and predictable anticoagulant effect with no need for routine monitoring. Although they are more expensive than warfarin, when the cost of monitoring INR and managing complications of VKA treatment is taken into account, they may represent a smaller financial burden. Some pharmacoeconomic studies have already been carried, but this is an area that will require further analysis over the next few years.

It must be remembered that there are still some important limitations to the new oral anticoagulants that need to be addressed. Probably the most important is the lack of a specific antidote or reversal agent. Although the drugs discussed here have significantly shorter half-lives when compared to warfarin, there are situations where immediate reversal may be required. There are some options available for these situations, but clinical experience is lacking and this is certainly an area where further research is required.

That no routine monitoring is required with the new oral anticoagulants is seen by most as a significant advantage, but it should be considered that this may have effects on patient compliance. Lack of compliance may be more difficult to identify or prove and, with the short half-life of these anticoagulants, a missed dose may be more significant than with warfarin. However, it could be argued that, even if this is the case, it hasn't been shown to be clinically significant in large randomized controlled trials.

Despite the many advantages of the new oral anticoagulants, warfarin is still likely to be used widely for many more years, especially in patients with mechanical valves, mitral stenosis, and cardiomyopathy. In time, new agents in these clinical situations may be evaluated, but until then warfarin and its unique clinical challenges will continue to be of relevance in everyday clinical practice.

REFERENCES

[1] Gomez K, Tuddenham EGD, McVey JH. Normal haemostasis. In: Hoffman R, Benz E, Shattil SJ, Furie B, Cohen H, editors. Hematology: basic principles and practice. 4th ed. Elsevier; 2005. p. 746–71.

[2] Adams RL, Bird RJ. Review article: coagulation cascade and therapeutics update: relevance to nephrology. Part 1: overview of coagulation, thrombophilias and history of anticoagulants. Nephrology (Carlton) 2009;14(5):462–70.

[3] Macfarlane RG. An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier. Nature 1964;2(202):498–9.

[4] Davie EW, Ratnoff OD. Waterfall sequence for intrinsic blood clotting. Science 1964;145(3638):1310–2.

[5] Monroe DM, Roberts HR, Hoffman M. Platelet procoagulant complex assembly in a tissue factor-initiated system. Br J Haematol 1994;88(2):364–71.

[6] Perez-Gomez F, Bover R. The new coagulation cascade and its possible influence on the delicate balance between thrombosis and hemorrhage. Rev Esp Cardiol 2007;60(12):1217–9.

[7] Hoffbrand AV, Catovsky D, Tuddenham EGD, Green AR. Postgraduate haematology. 6th ed. Wiley-Blackwell; 2005.

[8] Hoffman M, Monroe DM. Coagulation 2006: a modern view of hemostasis. Hematol Oncol Clin North Am 2007;21(1):1–11.

[9] Keeling D, Baglin T, Tait C, Watson H, Perry D, Baglin C, et al. Guidelines on oral anticoagulation with warfarin-fourth edition. Br J Haematol 2011;154(3):311–24.

[10] White RH. The epidemiology of venous thromboembolism. Circulation 2003;107(23 Suppl 1):I4–I8.

[11] Cushman M. Epidemiology and risk factors for venous thrombosis. Semin Hematol 2007;44(2):62–9.

[12] Kearon C, Crowther M, Hirsh J. Management of patients with hereditary hypercoagulable disorders. Annu Rev Med 2000;51:169–85.

[13] Linkins L, Kearon C. Venous thromboembolism. In: Key N,Makris M, O'Shaughnessy D, Lillicrap D, editors. Practical hemostasis and thrombosis. 2nd ed. Wiley-Blackwell; 2009. p. 135–46.

[14] National Institute for Health and Clinical Excellence. Venous thromboembolic diseases: the management of venous thromboembolic diseases and the role of thrombophilia testing. (Clinical guideline 144). NICE; 2012 <http://publications.nice.org.uk/venous-thromboembolic-diseases-the-management-of-venous-thromboembolic-diseases-and-the-role-of-cg144> [24.5.13 online source].

[15] Fuster V, Ryden LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, et al. ACC/AHA/ESC 2006 Guidelines for the management of patients with atrial fibrillation: a report of the american college of cardiology/american heart association task force on practice guidelines and the european society of cardiology committee for practice guidelines (writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation): developed in collaboration with the european heart rhythm association and the heart rhythm society. Circulation 2006;15(114(7)):e257–354.

[16] Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, et al. Guidelines for the management of atrial fibrillation: the task force for the management of atrial fibrillation of the european society of cardiology (ESC). Europace 2010;12(10):1360–420.

[17] Watson T, Shantsila E, Lip GY. Mechanisms of thrombogenesis in atrial fibrillation: Virchow's triad revisited. Lancet 2009;10(373(9658)):155–66.

[18] Wardrop D, Keeling D. The story of the discovery of heparin and warfarin. Br J Haematol 2008;141(6):757–63.

[19] Hirsh J, Warkentin TE, Shaughnessy SG, Anand SS, Halperin JL, Raschke R, et al. Heparin and low-molecular-weight heparin: mechanisms of action, pharmacokinetics, dosing, monitoring, efficacy, and safety. Chest 2001;119(Suppl 1):64S–94S.

[20] Garcia DA, Baglin TP, Weitz JI, Samama MM. Parenteral anticoagulants: antithrombotic therapy and prevention of thrombosis, 9th ed: American college of chest physicians evidence-based clinical practice guidelines. Chest 2012;141(2 Suppl):e24S–43S.

[21] Nutescu EA, Shapiro NL, Chevalier A, Amin AN. A pharmacologic overview of current and emerging anticoagulants. Cleve Clin J Med 2005;72(Suppl 1):S2–S6.

[22] Arepally GM, Ortel TL. Clinical practice. Heparin-induced thrombocytopenia. N Engl J Med 2006;24(355(8)):809–17.

[23] Denomme GA, Warkentin TE, Horsewood P, Sheppard JA, Warner MN, Kelton JG. Activation of platelets by sera containing IgG1 heparin-dependent antibodies: an explanation for the predominance of the Fc gammaRIIa "low responder" (his131) gene in patients with heparin-induced thrombocytopenia. J Lab Clin Med 1997;130(3):278–84.

[24] Ray KK, Francis S, Crossman DC. A potential pharmacogenomic strategy for anticoagulant treatment in non-ST elevation acute coronary syndromes: the role of interleukin-1 receptor antagonist genotype. J Thromb Haemost 2005;3(2):287–91.

[25] Yusuf S, Mehta SR, Chrolavicius S, Afzal R, Pogue J, Granger CB, et al. Comparison of fondaparinux and enoxaparin in acute coronary syndromes. N Engl J Med 2006;354(14):1464–76.

[26] Pirmohamed M. Warfarin: almost 60 years old and still causing problems. Br J Clin Pharmacol 2006;62(5):509–11.

[27] Ageno W, Gallus AS, Wittkowsky A, Crowther M, Hylek EM, Palareti G. Oral anticoagulant therapy: antithrombotic therapy and prevention of thrombosis, 9th ed: American college of chest physicians evidence-based clinical practice guidelines. Chest 2012;141(Suppl 2):e44S–88S.

[28] Holbrook AM, Pereira JA, Labiris R, McDonald H, Douketis JD, Crowther M, et al. Systematic overview of warfarin and its drug and food interactions. Arch Intern Med 2005;165(10):1095–106.

[29] Suvarna R, Pirmohamed M, Henderson L. Possible interaction between warfarin and cranberry juice. BMJ 2003;327(7429):1454.

[30] Kirkwood TB. Calibration of reference thromboplastins and standardisation of the prothrombin time ratio. Thromb Haemost 1983;49(3):238–44.

[31] Chiquette E, Amato MG, Bussey HI. Comparison of an anticoagulation clinic with usual medical care: anticoagulation control, patient outcomes, and health care costs. Arch Intern Med 1998;158(15):1641–7.

[32] Eriksson BI, Dahl OE, Rosencher N, Kurth AA, van Dijk CN, Frostick SP, et al. Dabigatran etexilate versus enoxaparin for prevention of venous thromboembolism after total hip replacement: a randomised, double-blind, non-inferiority trial. Lancet 2007;370(9591):949–56.

[33] Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 2009;361(12):1139–51.

[34] Schulman S, Kearon C, Kakkar AK, Schellong S, Eriksson H, Baanstra D, et al. Extended use of dabigatran, warfarin, or placebo in venous thromboembolism. N Engl J Med 2013;368(8):709–18.

[35] Hauel NH, Nar H, Priepke H, Ries U, Stassen JM, Wienen W. Structure-based design of novel potent nonpeptide thrombin inhibitors. J Med Chem 2002;45(9):1757–66.

[36] Stangier J, Stahle H, Rathgen K, Fuhr R. Pharmacokinetics and pharmacodynamics of the direct oral thrombin inhibitor dabigatran in healthy elderly subjects. Clin Pharmacokinet 2008;47(1):47–59.

[37] Garcia D, Libby E, Crowther MA. The new oral anticoagulants. Blood 2010;115(1):15–20.

[38] Ganetsky M, Babu KM, Salhanick SD, Brown RS, Boyer EW. Dabigatran: review of pharmacology and management of bleeding complications of this novel oral anticoagulant. J Med Toxicol 2011;7(4):281–7.

[39] Baglin T, Keeling D, Kitchen S. Effects on routine coagulation screens and assessment of anticoagulant intensity in patients taking oral dabigatran or rivaroxaban: guidance from the British Committee for Standards in Haematology. Br J Haematol 2012;159(4):427–9.

[40] Shepherd MF, Jacobsen JM, Rosborough TK. Argatroban therapy using enzymatic anti-factor IIa monitoring. Ann Pharmacother 2011;45(3):422–3.

[41] Fenyvesi T, Jorg I, Harenberg J. Monitoring of anticoagulant effects of direct thrombin inhibitors. Semin Thromb Hemost 2002;28(4):361–88.

[42] van Ryn J, Stangier J, Haertter S, Liesenfeld KH, Wienen W, Feuring M, et al. Dabigatran etexilate–a novel, reversible, oral direct thrombin inhibitor: interpretation of coagulation assays and reversal of anticoagulant activity. Thromb Haemost 2010;103(6):1116–27.

[43] van Ryn J, Ruehl D, Priepke H, Hauel N, Wienen W. Reversibility of the anticoagulant effect of high doses of the direct thrombin inhibitor dabigatran, by recombinant Factor VIIa or activated prothrombin complex concentrate (abstract 0370). Haematologica 2008;93:148.

[44] Eerenberg ES, Kamphuisen PW, Sijpkens MK, Meijers JC, Buller HR, Levi M. Reversal of rivaroxaban and dabigatran by prothrombin complex concentrate: a randomized, placebo-controlled, crossover study in healthy subjects. Circulation 2011;124(14):1573–9.

[45] Pare G, Eriksson N, Lehr T, Connolly S, Eikelboom J, Ezekowitz MD, et al. Genetic determinants of dabigatran plasma levels and their relation to bleeding. Circulation 2013;127(13):1404–12.

[46] Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenomics J 2008;8(3):186–95.

[47] Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med 2011;365(10):883–91.

[48] Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, et al. Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med 2011;365(11):981–92.

[49] Agnelli G, Buller HR, Cohen A, Curto M, Gallus AS, Johnson M, et al. Apixaban for extended treatment of venous thromboembolism. N Engl J Med 2013;368(8):699–708.

[50] Buller HR, Prins MH, Lensin AW, Decousus H, Jacobson BF, Minar E, et al. Oral rivaroxaban for the treatment of symptomatic pulmonary embolism. N Engl J Med 2012;366(14):1287–97.

[51] Cabral KP. Pharmacology of the new target-specific oral anticoagulants. J Thromb Thrombolysis 2013;36(2):133–40.

[52] Alquwaizani M, Buckley L, Adams C, Fanikos J. Anticoagulants: a review of the pharmacology, dosing, and complications. Curr Emerg Hosp Med Rep 2013;1(2):83–97.

[53] Harder S, Graff J. Novel oral anticoagulants: clinical pharmacology, indications and practical considerations. Eur J Clin Pharmacol 2013;69(9):1617–33.

[54] Adam SS, McDuffie JR, Ortel TL, Williams Jr JW. Comparative effectiveness of warfarin and new oral anticoagulants for the management of atrial fibrillation and venous thromboembolism: a systematic review. Ann Intern Med 2012;157(11):796–807.

Chapter 24

Pharmacogenomics of Warfarin

Ann K. Daly
Institute of Cellular Medicine, Newcastle University, UK

Chapter Outline

24.1 Overview 497
24.2 Pharmacology of Warfarin 498
24.3 Nongenetic Factors Affecting Warfarin Dose Requirement 499
24.4 Genes with Major Effects on Warfarin Dose Requirement 500
24.4.1 VKORC1: The Gene Encoding the Warfarin Target 500
24.4.2 CYP2C9: The Main Pharmacogenomic Factor Affecting Warfarin Pharmacokinetics 501
24.5 Additional Pharmacogenomic Factors Affecting Warfarin Dosage 502
24.5.1 CYP4F2 502
24.5.2 Calumenin 502
24.6 Application of Pharmacogenomics to Warfarin Prescription 503
24.7 Conclusion and Future Research Directions 504
References 504

24.1 OVERVIEW

Warfarin is an anticoagulant that was synthesized originally at the University of Wisconsin. Its name was derived from Wisconsin Alumni Research Foundation (*warf*) with *arin* taken from "heparin," another anticoagulant. It was used first as a rat poison but has been prescribed for the treatment of thromboembolic disorders since the 1950s. Warfarin remains one of the most widely prescribed drugs in modern medicine, but, unusually, dosage needs to be individualized for each patient. This normally involves starting with a standard dose, measuring the coagulation rate (international normalized ratio (INR)) after two days and then adjusting the dose by a computer algorithm until the required INR, usually between 2 and 3, is obtained. Overcoagulation places the patient at risk of potentially fatal hemorrhage whereas undercoagulation means that the patient does not benefit from the drug.

For these reasons, improving protocols for initiation of warfarin treatment is important. Most patients are treated successfully with warfarin using current protocols, but some individuals, including around 10% of Europeans and varying numbers from certain other ethnic groups, require a particularly low dose (1.5 mg/day or less). These patients may be at risk of too high a level of anticoagulation that could result in serious bleeding, especially during the initial treatment period [1].

In support of this, several studies have shown that warfarin is frequently linked to emergency hospital admissions relating to adverse drug reactions [2,3]. In addition to some individuals being unusually sensitive to warfarin and needing a low dose, a smaller group of individuals (no more than approximately 1% of the population) are "warfarin resistant." These patients need an unusually high dose of warfarin (more than 15 mg per day) to achieve the required level of anticoagulation; it may be that anticoagulation cannot be achieved at all. Alternatives to warfarin have been developed recently, but warfarin continues to be prescribed widely for newly diagnosed patients suffering from various thromboembolic disorders, including deep vein thrombosis and atrial fibrillation. Large numbers of patients continue to take warfarin on a long-term basis, mostly without serious problems.

Warfarin has a chemical structure homologous to coumarin and is often referred to as a coumarin anticoagulant.

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00024-4
© 2014 Elsevier Inc. All rights reserved.

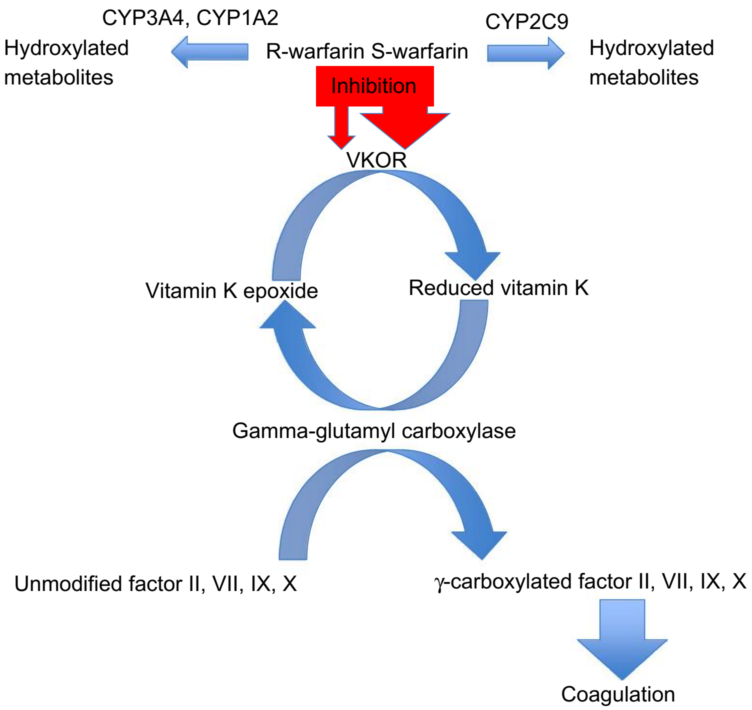


FIGURE 24.1 **The vitamin K cycle.** Warfarin targets VKOR, which regenerates reduced vitamin K after vitamin K epoxide is formed during the gamma carboxylase reaction.

Other coumarin anticoagulants (including acenocoumarol and phenprocoumon) are used as anticoagulants in preference to warfarin in certain European countries. Both drugs have considerable chemical homology to warfarin and overlap with it on pharmacogenomic factors affecting response. The genetic factors discussed in this chapter generally apply equally to these drugs, although there are some minor differences.

It is now very well established from a large number of studies worldwide that pharmacogenomic factors significantly affect individual warfarin dose requirement, although other nongenetic factors are also important. This is one of the best examples of a clinically relevant pharmacogenomic association, and attempts to translate genetic testing to the clinic prior to treatment are still ongoing. This chapter considers in detail a range of genetic polymorphisms in several genes that have been well validated as affecting warfarin dose requirement and also some additional genetic associations where the validation data are less complete, together with nongenetic factors. Progress to clinical translation is also considered.

24.2 PHARMACOLOGY OF WARFARIN

Some background on warfarin pharmacokinetics and pharmacodynamics is helpful in understanding the pharmacogenomic factors affecting response. Warfarin targets the vitamin K pathway, which plays a key role in the blood coagulation cascade. The clotting factors II, VII, IX, and X require post-translational modification of glutamic acid residues by gamma carboxylation prior to activation by proteolysis during the coagulation cascade. In the absence of this modification, the clotting factor is not recognized by the proteolysis complex. The carboxyl group is provided by CO_2 in a vitamin K-dependent reaction by the enzyme gamma-glutamyl carboxylase (see Figure 24.1). During this reaction, reduced vitamin K is converted to its oxidized form, vitamin K epoxide.

For further clotting reactions to occur, it is essential that the vitamin K epoxide is converted back to reduced vitamin K, because the latter is not stored in the body and dietary vitamin K is generally limiting. The enzyme vitamin K epoxide reductase (VKOR) catalyzes the conversion to reduced vitamin K, and warfarin acts by inhibiting VKOR by noncompetitive (also known as mixed) inhibition, whereby the inhibitor interacts with both the enzyme and the enzyme-substrate complex [4]. The gene encoding VKOR, termed vitamin K epoxide reductase complex 1 (VKORC1) was identified more recently than the majority of human genes.

Warfarin is a racemic drug with R and S enantiomers due to an asymmetric carbon at position 9 (see Figure 24.2). The S enantiomer shows a 3–5-fold higher inhibitory effect on VKOR compared with the R enantiomer. However, the

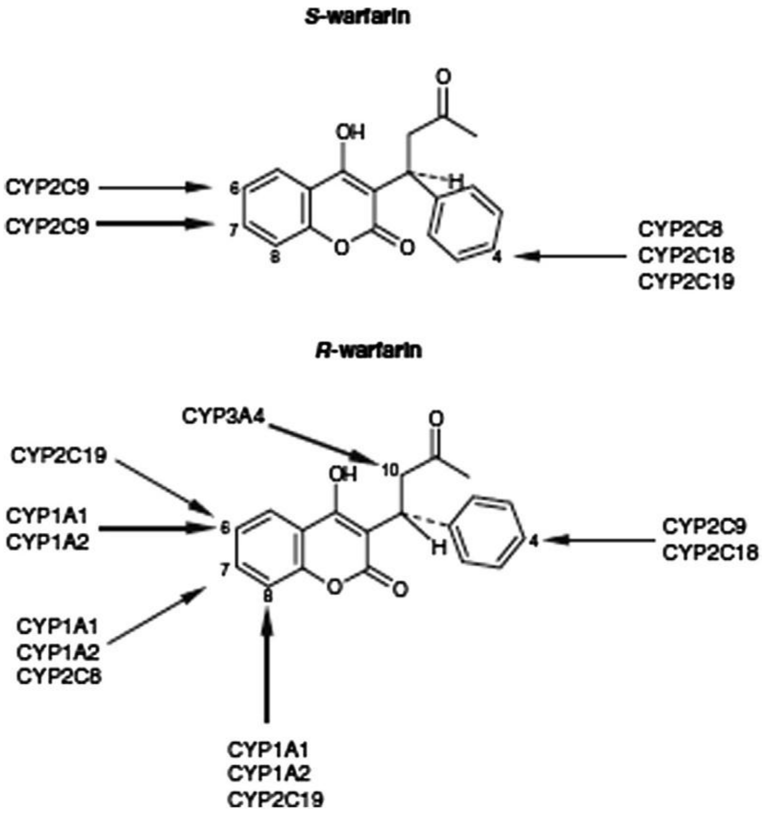


FIGURE 24.2 **S and R warfarin hydroxylation by cytochrome P450 enzymes.** Major metabolic routes are indicated by thick arrows; minor routes by thin arrows. *Source: Reprinted from Kaminsky LS, Zhang ZY*[7]*. Reproduced with permission.*

average plasma concentration of the R enantiomer is approximately two-fold higher than that of the S enantiomer, mainly because of differences in their metabolism [5]. S warfarin appears to account for 60–70% of the overall anticoagulation response, with the R enantiomer accounting for 30–40%. In the case of the S enantiomer, the major oxidative metabolite is the 7-hydroxy form, with some 6-hydroxy and 4′-hydroxy metabolites also formed at approximately equal concentrations. CYP2C9 is the main cytochrome P450 responsible for the 6- and 7-hydroxylation pathways, but at least two additional cytochromes P450, CYP2C8 and CYP2C19, can form the 4′-hydroxy metabolite [6,7]. R warfarin, on the other hand, shows more complex oxidative metabolism. The major metabolites include 6-, 8-, and 10-hydroxylated forms; 7′- and 4′-metabolites can also occur [7]. In the liver, the 6 and 8 metabolites are formed predominantly by CYP1A2. The 10-hydroxy metabolite is formed by CYP3A4—the cytochrome P450 expressed at the highest levels in most human livers. From kinetic data, formation of 10-hydroxy R warfarin by CYP3A4 is likely to be the most important oxidative metabolic route for this enantiomer.

In addition to oxidative metabolism by the cytochromes P450, the ketone group on the R warfarin side chain undergoes reduction by the carbonyl reductase enzyme family. After the phase I pathways of oxidation and reduction, the hydroxylated metabolites undergo further phase II metabolism by several UDP-glucuronosyltransferase isoforms prior to urinary excretion [8].

24.3 NONGENETIC FACTORS AFFECTING WARFARIN DOSE REQUIREMENT

Before considering genetic factors relevant to individual patient dose requirement with warfarin, it is helpful to consider the effect of nongenetic factors that are also important. One of the best studied factors is patient age. It is very well established that increasing age results in a lower dose requirement for warfarin [9]. The underlying mechanism is still not completely clear, but several independent studies suggest that it is not due to decreased warfarin metabolism [10]. Instead, it may be due to decreased hepatic blood flow slowing the delivery of warfarin to the liver [11,12], or it possibly may reflect poorer vitamin K absorption or intake in the elderly [13]. In general, a patient aged 75 years or older needs less than half the daily warfarin dose to achieve the same level of anticoagulation than does a patient aged 35 years or less [14]. When this is translated into the influence of age on warfarin dose, various estimates suggest that around 10% of variability in dose requirement is due to patient age [15,16].

Several studies have shown that parameters concerned with body size—that is, weight, height, or body mass index—also affect the warfarin dose necessary to achieve the required coagulation level [15,17,18]. A higher dose is needed where one or more of these factors is larger than average. The contribution from these parameters to dose variability is typically around 5%, so it is not as large as the variability associated with age.

Since warfarin affects the metabolism of vitamin K within the body, its action can be affected by dietary vitamin K intake. Though the actual contribution of vitamin K levels to individual dose requirements has not been estimated directly and is not taken into account when warfarin treatment is started, several studies have shown that patients who have low dietary vitamin K intake or whose intake varies have more problems maintaining a stable INR when undergoing treatment with warfarin than do patients with a higher intake that does not vary over time [19,20]. The possibility of giving patients vitamin K supplements when taking warfarin is currently under investigation. The warfarin dose needed by patients taking these supplements is higher, but better INR stability may be achieved. Two recent small-scale clinical trials have shown that this approach is feasible and safe, but further investigation in larger studies is needed to fully assess its usefulness [21,22].

Many patients undergoing treatment with warfarin are also taking additional prescribed drugs. These may interact with warfarin and, depending on the precise effect, either lower or raise the dose requirement. Not all prescribed drugs interact with warfarin, but a relatively large number of interactions have been reported [23]. The majority of drug interactions involve inhibition of warfarin metabolism or a decrease in its binding to plasma proteins, resulting in a lowering of dose requirement. Examples of drugs that have this effect include amiodarone, statins, and azole antifungals. However, some drugs are P450 inducers and may cause increased warfarin metabolism, resulting in a need to increase the dose. Examples of inducers include carbamazepine, phenytoin, and rifampicin. In addition to prescribed drugs, herbal medicines such as St. John's wort may also interact with warfarin. St. John's wort is an inducer that increases warfarin metabolism, but other herbal medicines may inhibit it. Tobacco smoking also interacts with warfarin, probably because of the cytochrome P450 CYP1A2, which metabolizes R warfarin being induced by components of tobacco smoke such as polycyclic aromatic hydrocarbons.

An additional factor that is known to affect INR stability is patient nonadherence [24,25]. As with any drug, the patient needs to be advised of the importance of adherence to the prescribed warfarin regimen.

24.4 GENES WITH MAJOR EFFECTS ON WARFARIN DOSE REQUIREMENT

Current data suggest that up to 40% of variability in individual warfarin dose requirements can be explained by genetic polymorphisms. Studies involving both candidate gene approaches and genome-wide association have demonstrated that two genes are particularly important. The VKORC1 and CYP2C9 genotypes together appear to account for the majority of variability, but a few other genes may make a more minor contribution. In addition, there is ethnic variation in the contribution of individual genes and the alleles involved.

24.4.1 VKORC1: The Gene Encoding the Warfarin Target

Until recently, the properties of the VKOR enzyme, especially the human enzyme, were poorly understood. The enzyme had not been purified, and knowledge of its properties was limited. In the original draft of the human genome sequence, no gene coding for this enzyme was shown because the gene had not been cloned and sequenced. However, because warfarin was used widely as a rat poison, data on rat VKOR was available, including the fact that some rats became resistant as a result of altered VKOR.

The genes encoding this protein in rats and mice were localized on the respective genome maps, and genome sequences for these species subsequently became available [26]. In parallel, progress on understanding human VKOR was made by investigations on two families in which several children were affected by bleeding soon after birth. Administration of vitamin K resulted in high levels of vitamin K epoxide, suggesting a defect in VKOR. Linkage analysis enabled the genetic defect to be localized to a region of chromosome 16 within a gene that appeared to code for VKOR and was termed VKORC1 [27].

This initial study also sequenced DNA from warfarin-resistant individuals and found that a majority of these were positive for nonsynonymous mutations in VKORC1. An independent study using siRNA to knock down genes encoding transmembrane proteins apparently associated with warfarin resistance also showed that the gene responsible for this trait was VKORC1 [28].

Subsequent studies on populations of patients undergoing warfarin treatment found that coding region polymorphisms in VKORC1 were rare, but identified upstream and intronic polymorphisms that correlated well with dose requirements. A number of different haplotypes were identified, but the general conclusion was that, in Europeans, the genotype for a single noncoding region polymorphism (either G-1639A or C1173T, which are in complete linkage disequilibrium) accounted for ~25% of variability in warfarin dose requirement, with the minor alleles associated with a lower dose requirement [16,29,30]. This was also the case in East Asians, but the minor allele occurred considerably more frequently than in Europeans, which explains the well-established finding of an average lower anticoagulant dose requirement in this geographic region [31,32].

Current data suggests that G-1639A affects VKOR transcription by the −1639 G allele associating preferentially with active chromatin compared with −1639A [33]. In Africans, the minor allele frequency for G-1639A/C1173T is considerably lower than in Europeans, and there are also differences in the VKORC1 haplotypes seen. The genotype for G-1639A or C1173T also predicts warfarin dose requirement in both Africans and African-Americans, but the proportion of variability in dose requirement in these groups explained by these genotypes is significantly lower (5−10%) compared with Europeans (20−30%), partly because both SNPs are less frequent in Africans.

Genotyping for additional common SNPs to detect a wider range of VKORC1 haplotypes in a study of African-Americans did not lead to an improved ability to predict variation [34]. However, a smaller study of African-Americans suggested that inclusion of a second VKORC1 SNP in addition to G-1639A improved prediction of warfarin dose [35]. Another recent study involving resequencing of VKORC1 in African-Americans found that genotyping for a common novel SNP situated more than 8 kb upstream of the VKORC1 translation start site (associated with an increased dose requirement), in addition to C1173T, increased the ability to predict warfarin dose requirement by greater than 3% [36].

Two additional VKORC1 SNPs, rs7200749 and rs7294, have been shown to be associated with increased warfarin dose requirement in black South Africans [37]. These SNPs were included in the earlier study on African-Americans, where no significant effect on warfarin dose was seen [34], but it is possible that the differences between the studies are due to a higher frequency of these SNPs among black South Africans.

Nonsynonymous mutations in VKORC1 explain many cases of warfarin resistance either where patients require an unusually high dose of the drug to achieve adequate anticoagulation, or it is not possible to achieve the required level of anticoagulation at all. A number of both isolated cases and family cases, where more than one family member was warfarin-resistant, were investigated when VKORC1 was identified [27]; they were found to be positive for such mutations. These findings have been confirmed subsequently by others.

It appears that amino acid substitutions in certain regions of VKOR result in conformational changes that prevent binding of the coumarin anticoagulant. Some of these mutations have been seen only in single patients, but others are more common. For example, Asp36Tyr was found to occur at a frequency of 4% in Ashkenazi Jews and at 15% in Ethiopian Jews [38]. Another mutation, Val66Met, is a relatively common cause of warfarin resistance in Europeans [27,39], but its overall effect at a population level is likely to be negligible.

Overall, current data suggest that genotyping for either G-1639A or C1173T is highly informative for predicting warfarin dose requirement, though the percentage variability depends on allele frequency in the individual population. Genotyping for Asp36Tyr may also be important in predicting dose required in some populations [40].

24.4.2 CYP2C9: The Main Pharmacogenomic Factor Affecting Warfarin Pharmacokinetics

CYP2C9 is a major cytochrome P450 isoform, both based on being a relatively abundant P450 in the liver and in terms of its overall contribution to P450-mediated drug oxidation. Probably the best studied CYP2C9 substrate is S warfarin, which undergoes 7-hydroxylation by CYP2C9. As discussed, S warfarin is also subject to metabolism at other positions by other CYP isoforms [6]. In addition, a different set of P450 isoforms metabolize R warfarin. Though the relatively complex metabolic profile of both warfarin enantiomers might suggest that a range of cytochrome P450 polymorphisms can affect metabolism and therefore be relevant to dose requirement, only the CYP2C9 genotype appears relevant to dose requirement based on a large number of published studies.

Most published studies have focused on the common variant alleles *CYP2C9*2* and **3*. These alleles are associated with amino acid substitutions that affect catalytic activity with S warfarin [41,42]. Initial pharmacogenetic studies showing a significant influence by the CYP2C9 genotype on dose requirement involved studies on patients known to require an unusually low dose of warfarin [1,43], but these findings were later confirmed in several larger studies involving patients of European ethnicity taking a wider range of warfarin doses [44–47]. Genome-wide association studies on European populations have also confirmed that *CYP2C9*2* and **3* are important contributors to warfarin dose requirement [48,49].

The overall contribution of the CYP2C9 genotype to warfarin dose requirement among Europeans has been estimated in a number of studies to be 6–19% [15,44,50]. Recent detailed studies have investigated genetic contributions to a variety of end points relevant to warfarin response in addition to dose requirement. From these, evidence has emerged that points to CYP2C9 as the most important genetic contributor to initial anticoagulant control [51,52], although not to stable anticoagulation.

Despite a large number of studies showing important roles for the *CYP2C9*2* and **3* alleles as determinants of warfarin dose requirement in Europeans, there have been fewer studies on other ethnic groups that show lower frequencies for these alleles but may be positive for other CYP2C9 variants that are also associated with decreased activity. For example, in African populations, the CYP2C9 allele distributions are very different from those described for Europeans. *CYP2C9*2* and **3* are not usually seen, but additional alleles associated with decreased enzyme activity, particularly *CYP2C9*8*, are common [37,53].

Other variant alleles, including *CYP2C9*5*, **6*, and **11*, have also been demonstrated to contribute to warfarin dose requirements in several studies on populations from various African countries. A similar range of CYP2C9 alleles is found among African-Americans from both North and South America, although *CYP2C9*2* and **3* are also relatively common among African-Americans [54]. In addition to the variants in the coding sequence mentioned, an intronic CYP2C9 polymorphism common in African-Americans was reported to be associated with an increased warfarin dose requirement [36], but this has not been confirmed in another independent study [55]. *CYP2C9*2* and *CYP2C9*3* are both detectable in South Asian populations although at frequencies lower than those seen in Europeans. Recent data suggest that *CYP2C9*8* occurs in South Indians [56].

The worldwide findings on the CYP2C9 genotype and warfarin dose requirement suggest that any algorithm for predicting warfarin dose should take account of the genotype for *CYP2C9*8* in addition to that for *CYP2C9*2* and **3*. There is more limited evidence that *CYP2C9*5*, **11* (which is seen at a low frequency among both Europeans and Africans), and **13* should also be considered (see Daly [57]).

24.5 ADDITIONAL PHARMACOGENOMIC FACTORS AFFECTING WARFARIN DOSAGE

The majority of published studies on pharmacogenomic factors affecting warfarin dosage agree that the VKORC1 and CYP2C9 genotypes are the important factors, with only minor contributions from other genes. This issue has been discussed in a recent detailed review article [57]; however, two of the genes, CYP4F2 and calumenin (CALU), appear to have a significant though small effect on warfarin dose requirement based on findings from a number of published studies, so these are discussed in more detail here.

24.5.1 CYP4F2

CYP4F2 is a cytochrome P450 isoform that plays a role in the metabolism of fatty acids and vitamin E [58,59]. A gene chip that genotypes for a large number of polymorphisms relevant to drug disposition, but does not cover the entire genome, provided novel data showing that the CYP4F2 genotype for the nonsynonymous polymorphism Val433Met (*CYP4F2*3* allele) contributes to warfarin dose requirement [60].

*CYP4F2*3* was found to be associated with a higher warfarin dose requirement. After its apparent relevance to warfarin treatment was reported, it was demonstrated that CYP4F2 also contributes to vitamin K1 oxidation, with the nonsynonymous polymorphism resulting in decreased activity *in vitro* [61]. A number of further studies, including several genome-wide association studies, have confirmed a contribution by CYP4F2 to warfarin dose requirement more generally [48,49]. Though the overall contribution of CYP4F2 to dose requirement appears relatively low (1–5%), *CYP4F2*3* is found at a fairly uniform frequency worldwide, so including it in an algorithm to predict warfarin dose may therefore be of value.

24.5.2 Calumenin

Calumenin is a member of the CREC family of Ca^{2+} binding proteins that serve as chaperones in certain secretory pathways [62]. Calumenin is expressed in a range of tissues, including heart, placenta, and skeletal muscle, and at a low level in the liver [63]. The protein is localized in the endoplasmic reticulum. A proposed mechanism of warfarin resistance in rats involves an interaction between VKOR and calumenin, with calumenin binding to VKOR to inhibit warfarin binding. It has also been demonstrated that livers of warfarin-resistant rats overexpress calumenin [64].

A further study on warfarin-resistant rats demonstrated that calumenin also inhibits gamma-carboxylase and therefore placed calumenin centrally in the gamma-carboxylation pathway [65]. There are a few reports on a possible contribution by the CALU genotype to warfarin dose requirement. In a study of the effect of known polymorphisms, in a range of candidate genes, on warfarin dose requirement in an Israeli population, a single patient was homozygous for a G11A allele in calumenin and required an unusually high warfarin dose of 20 mg/day [66]. The prevalence of heterozygotes for G11A in the upper tenth dose percentile was also significantly higher. More recently, a study of African-Americans that included sequencing of the CALU gene found that carriage of a noncoding SNP, which occurred at an allele frequency of 0.14 in this group, was associated with a high dose of warfarin (upper tenth percentile); it also found that this explained 5.7% of dose requirement variation [67].

In addition to CALU's positive associations with warfarin dose, a number of reports indicate no association of these genotypes with warfarin dose requirement or response. Calumenin is a good example of a biologically plausible candidate gene that requires further study to determine if it should be included in a warfarin dose prediction algorithm.

24.6 APPLICATION OF PHARMACOGENOMICS TO WARFARIN PRESCRIPTION

Most studies on genetic factors affecting warfarin response have focused on the relationship between genotype and stable dose requirement. In one of the earliest studies linking genotype and warfarin dose requirement, it was demonstrated that patients on a low, stable dose of warfarin are more likely to develop serious bleeds and high INR levels, which suggests that safety can be improved if genetic information is available at treatment initiation [1]. Subsequent detailed prospective studies have related both the CYP2C9 and the VKORC1 genotype to a range of clinical parameters from the initiation of warfarin treatment [51,52,54,68,69].

All such studies confirmed the significant contributions of CYP2C9 and VKORC1 as predictors of stable warfarin dose requirement. One study found CYP2C9 to be the most important gene determining initial anticoagulant control, and VKORC1 to be important for stable anticoagulation [51]. However, the largest study up to now found that homozygosity for CYP2C9 and VKORC1 variant alleles increased the risk of over-anticoagulation during the initiation period [52]. No effect for CYP2C9 was reported in a recent European study, but homozygosity for the variant VKORC1 allele was associated with both a shorter time to stable INR and a higher number of raised INR values (which increase the risk of bleeding) during the first month [69]. In a study involving both European and African-Americans, VKORC1 was revealed to be a significant predictor of overcoagulation during initiation in Europeans only, with a borderline significant effect for CYP2C9 but no significant effect by genotype on overcoagulation in African-Americans [54]. There is some disagreement between findings, but the various prospective studies indicate that further investigation of genotyping using randomized clinical trials may be worthwhile.

Clinical use of genotyping during warfarin treatment requires the development of suitable dosing algorithms to enable the initiation dose to be calculated. Since, as discussed previously, it is also well established that other factors, including patient age and size, and concurrent drugs affect dose requirement, these factors all need to be considered in the development of personalized dosing. The earliest dosing algorithms incorporating genetic factors were for warfarin in relation to the CYP2C9 genotype [15,44]. Once VKORC1 was also shown to be of relevance to dosing, algorithms incorporating this second genetic factor were developed [16,70–76]. Each of these algorithms was developed and tested on a relatively small patient group, though overall they were quite similar.

To improve prediction of dose requirement and to cover a wider range of ethnicities, a project involving a large number of researchers worldwide—the International Warfarin Pharmacogenetics Consortium (IWPC)—used clinical and genetic data on 4043 patients treated with warfarin to develop a more definitive dosing algorithm, which was then tested on a replication cohort of 1009 patients [18]. This was an important step forward in developing a clinically useful algorithm, but limitations included an overrepresentation of European patients and the fact that the only genotypes considered were those for common variant alleles in CYP2C9 and VKORC1. It was also uncertain exactly how an algorithm predicting a stable dose should be implemented clinically. Further refinements of the IWPC algorithm have subsequently been developed.

One refinement introduces a "dose revision algorithm," which involves initiation of treatment based on the IWPC algorithm-calculated dose followed by use of a new algorithm which incorporates genetic factors, INR value, and clinical factors on day 4 or 5 of treatment [77]. A second uses an "initiation dose" algorithm derived from the IWPC algorithm but introducing an additional term to increase the dose on days 1–3 of treatment [78]. Few currently available warfarin dosing algorithms include genetic factors other than the common CYP2C9 and VKORC1 variants, but a modification of the IWPC algorithm to include CYP4F2 genotype and a range of rarer CYP2C9 alleles was said to improve stable dose prediction [70].

Though the IWPC algorithm was developed to cover a wide range of ethnic groups, there are some recent examples of warfarin dosing algorithms developed specifically for certain ethnic groups [79,80]. Children occasionally require treatment with warfarin. Two recent studies suggest that the algorithms developed for adults are not suitable for use in children, and have developed alternatives [81,82].

A number of small clinical trials of varying designs, including some involving a randomized approach, have examined the value of using genotype and other patient-related factors to set initial warfarin dose [73,83–87]. Most of these report no advantage for a genotype-determined dose, but they may not have had adequate power to detect all genetic effects. One study focused on CYP2C9 genotypes only and did find improved time to stable INR and a higher percentage of time within therapeutic range in the group with the genotype-guided dose compared with controls receiving standard care [83]. Another obtained data showing that provision of information on genotype for both CYP2C9 and VKORC1 to the treating physician resulted in better outcomes for warfarin treatment, although genotype data was provided on average 32 days after the start of treatment, which is a serious limitation [84].

The only large randomized control trial up to the present involved 504 cases treated with a pharmacogenetic-guided dose and 1911 controls given standard dosing [88].

It showed that the pharmacogenetic-guided dosing group had a higher percentage of patients within therapeutic INR range at two time points, and fewer patients showing serious adverse events or very low or high INR values. Further randomized control trials of this type are in progress in Europe, the United States, and Singapore [89–91].

If the large clinical trials currently in progress report positive findings, more widespread translation of genotyping to the initiation of warfarin dosing seems likely. Some guidelines in relation to warfarin treatment genotyping have already been provided [92]. These guidelines are particularly aimed at the U.S. healthcare system, but should be relevant elsewhere. In particular, they emphasize that the likely value of genotyping comes early in the course of therapy, either ideally before the start or, if this is not possible, within the first few days.

Obtaining genotypes rapidly remains challenging, but some progress has been made in developing point-of-care systems. These include methods specifically geared toward use of warfarin [93] and methods to genotype for alternative pharmacogenomic polymorphisms which could be adapted for warfarin [94]. It also seems possible that, in the future, obtaining this type of genetic data might be a routine part of healthcare. Thus, pharmacogenomic information would be already available if needed [95].

24.7 CONCLUSION AND FUTURE RESEARCH DIRECTIONS

As discussed in this chapter, there is widespread agreement on a significant relationship between warfarin dose requirement and CYP2C9 and VKORC1 genotypes. However, there is much less agreement on the value of using genotyping to predict dose requirement. Even if the various clinical trials confirm that a personalized medicine approach involving genotyping leads to improved outcomes in the prescription of warfarin, it is possible that the availability of alternative drugs, including direct thrombin inhibitors such as dabigatran, may make warfarin treatment obsolete and no longer attractive for implementation of personalized medicine.

Some recent modeling data on dabigatran suggests that, if time within target INR can be improved for warfarin by incorporation of genotyping, dabigatran will likely be less cost-effective than continuing warfarin use [96]. Although dabigatran and other similar drugs do not require routine INR monitoring, there is still a risk that bleeding may occur. A recent genome-wide association study on dabigatran pharmacogenomics [97] found decreased bleeding risk in those positive for an intronic SNP in the carboxyesterase gene *CES1*, which encodes an enzyme that metabolizes dabigatran. It was suggested that genotyping for this SNP might be beneficial in individualizing dabigatran dose to increase safety [97]. This slightly unexpected finding may help in demonstrating cost-effectiveness for warfarin prescription combined with genotyping.

Key Concepts

- Warfarin is used widely as an anticoagulant drug and is unusual in that dosage is individualized for each patient by measurement of response.
- The dose required varies widely among patients.
- Factors affecting dose requirement include patient age, size, dietary vitamin K, concomitant drugs, and genotype for several genetic polymorphisms.
- The warfarin target enzyme, vitamin K epoxide reductase, is encoded by VKORC1. Common noncoding polymorphisms in VKORC1 affect warfarin dose requirement significantly and may account for 20–30% of variation in dose requirement.
- The more active warfarin form, S warfarin, is subject to metabolism by the cytochrome P450 CYP2C9. The gene encoding CYP2C9 has two common nonsynonymous polymorphisms that are associated with decreased activity. Patients positive for one or both of these require lower doses of warfarin. The CYP2C9 genotype accounts for 10–20% of variability in dose requirement.
- Other genetic factors may also contribute. The best studied examples are the genotype for CYP4F2 and calumenin, each of which may have a small effect on dose requirement.
- Randomized clinical trials to test the use of a dosing algorithm based on patient age, size, and VKORC1 and CYP2C9 genotypes to determine the initial dose of warfarin are in progress.

REFERENCES

[1] Aithal GP, Day CP, Kesteven PJL, Daly AK. Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet 1999;353(9154):717–9.

[2] Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18,820 patients. BMJ 2004;329(7456):15–19.

[3] Carrasco-Garrido P, de Andres LA, Barrera VH, de Miguel GA, Jimenez-Garcia R. Trends of adverse drug reactions related-hospitalizations in Spain (2001–2006). BMC Health Serv Res 2010;10:287.

[4] Bevans CG, Krettler C, Reinhart C, Tran H, Kossmann K, Watzka M, et al. Determination of the warfarin inhibition constant Ki for vitamin K 2,3-epoxide reductase complex subunit-1 (VKORC1) using an *in vitro* DTT-driven assay. Biochim Biophys Acta 2013;1830(8):4202–10.

[5] Takahashi H, Echizen H. Pharmacogenetics of warfarin elimination and its clinical implications. Clin Pharmacokinet 2001;40(8):587–603.

[6] Rettie AE, Korzekwa KR, Kunze KL, Lawrence RF, Eddy AC, Aoyama T, et al. Hydroxylation of warfarin by human cDNA-expressed cytochrome-P-450 – a role for P-4502c9 in the etiology of (S)-warfarin drug-interactions. Chem Res Toxicol 1992;5(1):54–9.

[7] Kaminsky LS, Zhang ZY. Human P450 metabolism of warfarin. Pharmacol Ther 1997;73(1):67–74.

[8] Zielinska A, Lichti CF, Bratton S, Mitchell NC, Gallus-Zawada A, Le VH, et al. Glucuronidation of monohydroxylated warfarin metabolites by human liver microsomes and human recombinant UDP-glucuronosyltransferases. J Pharmacol Exp Ther 2008;324(1):139–48.

[9] Gurwitz JH, Avorn J, Ross-Degnan D, Choodnovskiy I, Ansell J. Aging and the anticoagulant response to warfarin therapy. Ann Intern Med 1992;116(11):901–4.

[10] Routledge PA, Chapman PH, Davies DM, Rawlins MD. Pharmacokinetics and pharmacodynamics of warfarin at steady state. Br J Clin Pharmacol 1979;8(3):243–7.

[11] Woodhouse K, Wynne HA. Age-related changes in hepatic function. Implications for drug therapy. Drugs Aging 1992;2(3):243–55.

[12] Wynne HA, Cope LH, Mutch E, Rawlins MD, Woodhouse KW, James OF. The effect of age upon liver volume and apparent liver blood flow in healthy man. Hepatology 1989;9(2):297–301.

[13] Booth SL, Al Rajabi A. Determinants of vitamin K status in humans. Vitam Horm 2008;78:1–22.

[14] Fitzmaurice DA, Blann AD, Lip GY. Bleeding risks of antithrombotic therapy. BMJ 2002;325(7368):828–31.

[15] Kamali F, Khan TI, King BP, Frearson R, Kesteven P, Wood P, et al. Contribution of age, body size, and CYP2C9 genotype to anticoagulant response to warfarin. Clin Pharmacol Ther 2004;75(3):204–12.

[16] Sconce EA, Khan TI, Wynne HA, Avery P, Monkhouse L, King BP, et al. The impact of CYP2C9 and VKORC1 genetic polymorphism and patient characteristics upon warfarin dose requirements: proposal for a new dosing regimen. Blood 2005;106:2329–33.

[17] Absher RK, Moore ME, Parker MH. Patient-specific factors predictive of warfarin dosage requirements. Ann Pharmacother 2002;36(10):1512–7.

[18] Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MTM, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med 2009;360(8):753–64.

[19] Sconce E, Khan T, Mason J, Noble F, Wynne H, Kamali F. Patients with unstable control have a poorer dietary intake of vitamin K compared to patients with stable control of anticoagulation. Thromb Haemost 2005;93(5):872–5.

[20] Kim KH, Choi WS, Lee JH, Lee H, Yang DH, Chae SC. Relationship between dietary vitamin K intake and the stability of anticoagulation effect in patients taking long-term warfarin. Thromb Haemost 2010;104(4):755–9.

[21] Sconce E, Avery P, Wynne H, Kamali F. Vitamin K supplementation can improve stability of anticoagulation for patients with unexplained variability in response to warfarin. Blood 2007;109(6):2419–23.

[22] de Assis MC, Rabelo ER, Avila CW, Polanczyk CA, Rohde LE. Improved oral anticoagulation after a dietary vitamin k-guided strategy: a randomized controlled trial. Circulation 2009;120(12):1115–22. [3 p following 22]

[23] Holbrook AM, Pereira JA, Labiris R, McDonald H, Douketis JD, Crowther M, et al. Systematic overview of warfarin and its drug and food interactions. Arch Intern Med 2005;165(10):1095–106.

[24] Waterman AD, Milligan PE, Bayer L, Banet GA, Gatchel SK, Gage BF. Effect of warfarin nonadherence on control of the International Normalized Ratio. Am J Health Syst Pharm 2004;61(12):1258–64.

[25] Platt AB, Localio AR, Brensinger CM, Cruess DG, Christie JD, Gross R, et al. Can we predict daily adherence to warfarin? Results from the International Normalized Ratio Adherence and Genetics (IN-RANGE) study. Chest 2010;137(4):883–9.

[26] Kohn MH, Pelz HJ. A gene-anchored map position of the rat warfarin-resistance locus, Rw, and its orthologs in mice and humans. Blood 2000;96(5):1996–8.

[27] Rost S, Fregin A, Ivaskevicius V, Conzelmann E, Hortnagel K, Pelz HJ, et al. Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2. Nature 2004;427(6974):537–41.

[28] Li T, Chang CY, Jin DY, Lin PJ, Khvorova A, Stafford DW. Identification of the gene for vitamin K epoxide reductase. Nature 2004;427(6974):541–4.

[29] Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, McLeod HL, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N Engl J Med 2005;352(22):2285–93.

[30] D'Andrea G, D'Ambrosio RL, Di Perna P, Chetta M, Santacroce R, Brancaccio V, et al. A polymorphism in VKORC1 gene is associated with an inter-individual variability in the dose-anticoagulant effect of warfarin. Blood 2005;105(2):645–9.

[31] Yuan HY, Chen JJ, Lee MTM, Wung JC, Chen YF, Charng MJ, et al. A novel functional VKORC1 promoter polymorphism is associated with inter-individual and inter-ethnic differences in warfarin sensitivity. Hum Mol Genet 2005;14(13):1745–51.

[32] Veenstra DL, Blough DK, Higashi MK, Farin FM, Srinouanprachan S, Rieder MJ, et al. CYP2C9 haplotype structure in European American warfarin patients and association with clinical outcomes. Clin Pharmacol Ther 2005;77(5):353–64.

[33] Wang D, Chen H, Momary KM, Cavallari LH, Johnson JA, Sadee W. Regulatory polymorphism in vitamin K epoxide reductase complex subunit 1 (VKORC1) affects gene expression and warfarin dose requirement. Blood 2008;112(4):1013–21.

[34] Limdi NA, Wadelius M, Cavallari L, Eriksson N, Crawford DC, Lee MT, et al. Warfarin pharmacogenetics: a single VKORC1 polymorphism is predictive of dose across 3 racial groups. Blood 2010;115(18):3827–34.

[35] Schelleman H, Brensinger CM, Chen J, Finkelman BS, Rieder MJ, Kimmel SE. New genetic variant that might improve warfarin dose prediction in African Americans. Br J Clin Pharmacol 2010;70(3):393–9.

[36] Perera MA, Gamazon E, Cavallari LH, Patel SR, Poindexter S, Kittles RA, et al. The missing association: sequencing-based discovery of novel SNPs in VKORC1 and CYP2C9 that affect warfarin dose in African Americans. Clin Pharmacol Ther 2011;89(3):408–15.

[37] Mitchell C, Gregersen N, Krause A. Novel CYP2C9 and VKORC1 gene variants associated with warfarin dosage variability in the South African black population. Pharmacogenomics 2012;12(7):953–63.

[38] Loebstein R, Dvoskin I, Halkin H, Vecsler M, Lubetsky A, Rechavi G, et al. A coding VKORC1 Asp36Tyr polymorphism predisposes to warfarin resistance. Blood 2007;109(6):2477–80.

[39] Harrington DJ, Gorska R, Wheeler R, Davidson S, Murden S, Morse C, et al. Pharmacodynamic resistance to warfarin is associated with nucleotide substitutions in VKORC1. J Thromb Haemost 2008;6(10):1663–70.

[40] Kurnik D, Qasim H, Sominsky S, Lubetsky A, Markovits N, Li C, et al. Effect of the VKORC1 D36Y variant on warfarin dose requirement and pharmacogenetic dose prediction. Thromb Haemost 2012;108(4):781–8.

[41] Crespi CL, Miller VP. The R144C change in the CYP2C9*2 allele alters interaction of the cytochrome P450 with NADPH: cytochrome P450 oxidoreductase. Pharmacogenetics 1997;7(3):203–10.

[42] Haining RL, Hunter AP, Veronese ME, Trager WF, Rettie AE. Allelic variants of human cytochrome P450 2C9: baculovirus-mediated expression, purification, structural characterization, substrate stereoselectivity, and prochiral selectivity of the wild-type and I359L mutant forms. Arch Biochem Biophys 1996;333(2):447–58.

[43] Steward DJ, Haining RL, Henne KR, Davis G, Rushmore TH, Trager WF, et al. Genetic association between sensitivity to warfarin and expression of CYP2C9*3. Pharmacogenetics 1997;7(5):361–7.

[44] Gage BF, Eby C, Milligan PE, Banet GA, Duncan JR, McLeod HL. Use of pharmacogenetics and clinical factors to predict the maintenance dose of warfarin. Thromb Haemost 2004;91(1):87–94.

[45] Higashi MK, Veenstra DL, Kondo LML, Wittkowsky AK, Srinouanprachanh SL, Farin FM, et al. Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. JAMA 2002;287(13):1690–8.

[46] Taube J, Halsall D, Baglin T. Influence of cytochrome P-450CYP2C9 polymorphisms on warfarin sensitivity and risk of over-anticoagulation in patients on long-term treatment. Blood 2000;96(5):1816–9.

[47] Wadelius M, Sorlin K, Wallerman O, Karlsson J, Yue QY, Magnusson PK, et al. Warfarin sensitivity related to CYP2C9, CYP3A5, ABCB1 (MDR1) and other factors. Pharmacogenomics J 2004;4(1):40–8.

[48] Cooper GM, Johnson JA, Langaee TY, Feng H, Stanaway IB, Schwarz UI, et al. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. Blood 2008;112(4):1022–7.

[49] Takeuchi F, McGinnis R, Bourgeois S, Barnes C, Eriksson N, Soranzo N, et al. A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. PLoS Genet 2009;5(3):e1000433.

[50] Hillman MA, Wilke RA, Caldwell MD, Berg RL, Glurich I, Burmester JK. Relative impact of covariates in prescribing warfarin according to CYP2C9 genotype. Pharmacogenetics 2004;14(8):539–47.

[51] Jorgensen AL, Al-Zubiedi S, Zhang JE, Keniry A, Hanson A, Hughes DA, et al. Genetic and environmental factors determining clinical outcomes and cost of warfarin therapy: a prospective study. Pharmacogenet Genomics 2009;19(10):800–12.

[52] Wadelius M, Chen LY, Lindh JD, Eriksson N, Ghori MJ, Bumpstead S, et al. The largest prospective warfarin-treated cohort supports genetic forecasting. Blood 2009;113(4):784–92.

[53] Allabi AC, Gala JL, Horsmans Y, Babaoglu MO, Bozkurt A, Heusterspreute M, et al. Functional impact of CYP2C*5, CYP2C9*6, CYP2C9*8, and CYP2C9*11 *in vivo* among black Africans. Clin Pharmacol Ther 2004;76(2):113–8.

[54] Limdi NA, Arnett DK, Goldstein JA, Beasley TM, McGwin G, Adler BK, et al. Influence of CYP2C9 and VKORC1 on warfarin dose, anticoagulation attainment and maintenance among European-Americans and African-Americans. Pharmacogenomics 2008;9(5):511–26.

[55] Cavallari LH, Perera M, Wadelius M, Deloukas P, Taube G, Patel SR, et al. Association of the GGCX (CAA)16/17 repeat polymorphism with higher warfarin dose requirements in African Americans. Pharmacogenet Genomics 2012;22(2):152–8.

[56] Pavani A, Naushad SM, Mishra RC, Malempati AR, Pinjala R, Kumar TR, et al. Retrospective evidence for clinical validity of expanded genetic model in warfarin dose optimization in a South Indian population. Pharmacogenomics 2012;13(8):869–78.

[57] Daly AK. Optimal dosing of warfarin and other coumarin anticoagulants: the role of genetic polymorphisms. Arch Toxicol 2013;87(3):407–20.

[58] Sontag TJ, Parker RS. Influence of major structural features of tocopherols and tocotrienols on their omega-oxidation by tocopherol-omega-hydroxylase. J Lipid Res 2007;48(5):1090–8.

[59] Stec DE, Roman RJ, Flasch A, Rieder MJ. Functional polymorphism in human CYP4F2 decreases 20-HETE production. Physiol Genomics 2007;30(1):74–81.

[60] Caldwell MD, Awad T, Johnson JA, Gage BF, Falkowski M, Gardina P, et al. CYP4F2 genetic variant alters required warfarin dose. Blood 2008;111(8):4106–12.

[61] McDonald MG, Rieder MJ, Nakano M, Hsia CK, Rettie AE. CYP4F2 is a vitamin K1 oxidase: an explanation for altered warfarin dose in carriers of the V433M variant. Mol Pharmacol 2009;75(6):1337–46.

[62] Vorum H, Hager H, Christensen BM, Nielsen S, Honore B. Human calumenin localizes to the secretory pathway and is secreted to the medium. Exp Cell Res 1999;248(2):473–81.

[63] Yabe D, Taniwaki M, Nakamura T, Kanazawa N, Tashiro K, Honjo T. Human calumenin gene (CALU): cDNA isolation and chromosomal mapping to 7q32. Genomics 1998;49(2):331–3.

[64] Wallin R, Hutson SM, Cain D, Sweatt A, Sane DC. A molecular mechanism for genetic warfarin resistance in the rat. Faseb J 2001;15(11):U163–86.

[65] Wajih N, Sane DC, Hutson SM, Wallin R. The inhibitory effect of calumenin on the vitamin K-dependent gamma-carboxylation system. Characterization of the system in normal and warfarin-resistant rats. J Biol Chem 2004;279(24):25276–83.

[66] Vecsler M, Loebstein R, Almog S, Kurnik D, Goldman B, Halkin H, et al. Combined genetic profiles of components and regulators of the vitamin K-dependent gamma-carboxylation system affect individual sensitivity to warfarin. Thromb Haemost 2006;95(2):205–11.

[67] Voora D, Koboldt DC, King CR, Lenzini PA, Eby CS, Porche-Sorbet R, et al. A polymorphism in the VKORC1 regulator calumenin predicts higher warfarin dose requirements in African Americans. Clin Pharmacol Ther 2010;87(4):445–51.

[68] Limdi NA, Wiener H, Goldstein JA, Acton RT, Beasley TM. Influence of CYP2C9 and VKORC1 on warfarin response during initiation of therapy. Blood Cells Mol Dis 2009;43(1):119–28.

[69] Lund K, Gaffney D, Spooner R, Etherington AM, Tansey P, Tait RC. Polymorphisms in VKORC1 have more impact than CYP2C9 polymorphisms on early warfarin International Normalized Ratio control and bleeding rates. Br J Haematol 2012;158(2):256–61.

[70] Herman D, Peternel P, Stegnar M, Breskvar K, Dolzan V. The influence of sequence variations in factor VII, gamma-glutamyl carboxylase and vitamin K epoxide reductase complex genes on warfarin dose requirement. Thromb Haemost 2006;95(5):782–7.

[71] Tham LS, Goh BC, Nafziger A, Guo JY, Wang LZ, Soong R, et al. A warfarin-dosing model in Asians that uses single-nucleotide polymorphisms in vitamin K epoxide reductase complex and cytochrome P450 2C9. Clin Pharmacol Ther 2006;80(4):346–55.

[72] Caldwell MD, Berg RL, Zhang KQ, Glurich I, Schmelzer JR, Yale SH, et al. Evaluation of genetic factors for warfarin dose prediction. Clin Med Res 2007;5(1):8–16.

[73] Anderson JL, Horne BD, Stevens SM, Grove AS, Barton S, Nicholas ZP, et al. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. Circulation 2007;116(22):2563–70.

[74] Millican EA, Lenzini PA, Milligan PE, Grosso L, Eby C, Deych E, et al. Genetic-based dosing in orthopedic patients beginning warfarin therapy. Blood 2007;110(5):1511–5.

[75] Zhu Y, Shennan M, Reynolds KK, Johnson NA, Herrnberger MR, Valdes Jr R, et al. Estimation of warfarin maintenance dose based on VKORC1 (−1639 G>A) and CYP2C9 genotypes. Clin Chem 2007;53(7):1199–205.

[76] Wu AH, Wang P, Smith A, Haller C, Drake K, Linder M, et al. Dosing algorithm for warfarin using CYP2C9 and VKORC1 genotyping from a multi-ethnic population: comparison with other equations. Pharmacogenomics 2008;9(2):169–78.

[77] Lenzini P, Wadelius M, Kimmel S, Anderson JL, Jorgensen AL, Pirmohamed M, et al. Integration of genetic, clinical, and INR data to refine warfarin dosing. Clin Pharmacol Ther 2010;87(5):572–8.

[78] Avery PJ, Jorgensen A, Hamberg AK, Wadelius M, Pirmohamed M, Kamali F. A proposal for an individualized pharmacogenetics-based warfarin initiation dose regimen for patients commencing anticoagulation therapy. Clin Pharmacol Ther 2011;90(5):701–6.

[79] Cho HJ, On YK, Bang OY, Kim JW, Huh W, Ko JW, et al. Development and comparison of a warfarin-dosing algorithm for Korean patients with atrial fibrillation. Clin Ther 2011;33(10):1371–80.

[80] Wei M, Ye F, Xie D, Zhu Y, Zhu J, Tao Y, et al. A new algorithm to predict warfarin dose from polymorphisms of CYP4F2, CYP2C9 and VKORC1 and clinical variables: derivation in Han Chinese patients with non valvular atrial fibrillation. Thromb Haemost 2012;107(6):1083–91.

[81] Biss TT, Avery PJ, Brandao LR, Chalmers EA, Williams MD, Grainger JD, et al. VKORC1 and CYP2C9 genotype and patient characteristics explain a large proportion of the variability in warfarin dose requirement among children. Blood 2012;119(3):868–73.

[82] Moreau C, Bajolle F, Siguret V, Lasne D, Golmard JL, Elie C, et al. Vitamin K antagonists in children with heart disease: height and VKORC1 genotype are the main determinants of the warfarin dose requirement. Blood 2012;119(3):861–7.

[83] Caraco Y, Blotnick S, Muszkat M. CYP2C9 genotype-guided warfarin prescribing enhances the efficacy and safety of anticoagulation: a prospective randomized controlled study. Clin Pharmacol Ther 2008;83(3):460–70.

[84] Epstein RS, Moyer TP, Aubert RE, Dj OK, Xia F, Verbrugge RR, et al. Warfarin genotyping reduces hospitalization rates results from the MM-WES (Medco-Mayo warfarin effectiveness study). J Am Coll Cardiol 2010;55(25):2804–12.

[85] McMillin GA, Melis R, Wilson A, Strong MB, Wanner NA, Vinik RG, et al. Gene-based warfarin dosing compared with standard of care practices in an orthopedic surgery population: a prospective, parallel cohort study. Ther Drug Monit 2010;32(3):338–45.

[86] Burmester JK, Berg RL, Yale SH, Rottscheit CM, Glurich IE, Schmelzer JR, et al. A randomized controlled trial of genotype-based Coumadin initiation. Genet Med 2011;13(6):509–18.

[87] Gong IY, Tirona RG, Schwarz UI, Crown N, Dresser GK, Larue S, et al. Prospective evaluation of a pharmacogenetics-guided warfarin loading and maintenance dose regimen for initiation of therapy. Blood 2011;118(11):3163–71.

[88] Anderson JL, Horne BD, Stevens SM, Woller SC, Samuelson KM, Mansfield JW, et al. A randomized and clinical effectiveness trial comparing two pharmacogenetic algorithms and standard care for individualizing warfarin dosing (CoumaGen-II). Circulation 2012;125(16):1997–2005.

[89] van Schie RM, Wadelius MI, Kamali F, Daly AK, Manolopoulos VG, de Boer A, et al. Genotype-guided dosing of coumarin derivatives: the European pharmacogenetics of anticoagulant therapy (EU-PACT) trial design. Pharmacogenomics 2009;10(10):1687–95.

[90] French B, Joo J, Geller NL, Kimmel SE, Rosenberg Y, Anderson JL, et al. Statistical design of personalized medicine interventions: the Clarification of Optimal Anticoagulation through Genetics (COAG) trial. Trials 2010;11:108.

[91] Do EJ, Lenzini P, Eby CS, Bass AR, McMillin GA, Stevens SM, et al. Genetics informatics trial (GIFT) of warfarin to prevent deep vein thrombosis (DVT): rationale and study design. Pharmacogenomics J 2012;12(5):417–24.

[92] Johnson JA, Gong L, Whirl-Carrillo M, Gage BF, Scott SA, Stein CM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. Clin Pharmacol Ther 2011;90(4):625–9.

[93] Howard R, Leathart JBS, French DJ, Krishan E, Kohnke H, Wadelius M, et al. Genotyping for CYP2C9 and VKORC1 alleles by a novel point of care assay with HyBeacon (R) probes. Clin Chim Acta 2011;412(23–24):2063–9.

[94] Roberts JD, Wells GA, Le May MR, Labinaz M, Glover C, Froeschl M, et al. Point-of-care genetic testing for personalisation of antiplatelet treatment (RAPID GENE): a prospective, randomised, proof-of-concept trial. Lancet 2012;379(9827):1705–11.

[95] Gottesman O, Scott SA, Ellis SB, Overby CL, Ludtke A, Hulot JS, et al. The CLIPMERGE PGx program: clinical implementation of personalized medicine through electronic health records and genomics-pharmacogenomics. Clin Pharmacol Ther 2013;94(2):214–7.

[96] You JHS, Tsui KKN, Wong RSM, Cheng G. Cost-effectiveness of dabigatran versus genotype-guided management of warfarin therapy for stroke prevention in patients with atrial fibrillation. PLoS ONE 2012;7(6):e39640.

[97] Pare G, Eriksson N, Lehr T, Connolly S, Eikelboom J, Ezekowitz MD, et al. Genetic determinants of dabigatran plasma levels and their relation to bleeding. Circulation 2013;127(13):1404–12.

This page is intentionally left blank

Chapter 25

Pharmacogenomics of Clopidogrel

Betti Giusti, Claudia Saracini, Silvia Galora, and Rossella Marcucci
Department of Experimental and Clinical Medicine, University of Florence, Atherothrombotic Diseases Center, Careggi Hospital, Italy

Chapter Outline

25.1 Overview 509
  25.1.1 Platelets, Atherothrombotic Diseases, and Antiplatelet Therapy 509
  25.1.2 Clopidogrel Metabolism 511
  25.1.3 Alternative Thienopyridine-Class Chemical Compounds 513
25.2 Clopidogrel Response Variation 514
  25.2.1 High on-Clopidogrel Platelet Reactivity 514
  25.2.2 Nonpharmacogenetic Mechanisms of High on-Clopidogrel Platelet Reactivity 515
25.3 Genetic Dissection of Clopidogrel Response 518
  25.3.1 Genetic Determinants of Clopidogrel Response Variability 518
  25.3.2 Association of *CYP2C19*2* with High on-Clopidogrel Platelet Reactivity in Healthy Subjects 519
  25.3.3 Association of *CYP2C19*2* with High on-Clopidogrel Platelet Reactivity in High-Risk Vascular Patients 522
  25.3.4 Association of *CYP2C19*2* with Major Adverse Cardiovascular Events 523
  25.3.5 Other Investigated Genetic Variants 526
25.4 Current Status 527
  25.4.1 Screening, Stratification, and Personalization 527
  25.4.2 Pharmacogenetic Testing 529
25.5 Conclusions and Future Directions 535
Abbreviations 536
References 536

25.1 OVERVIEW

Platelets are the smallest elements of the blood; physiologically, they play a crucial role in the hemostatic process. Platelets interact with numerous cells, playing a key role in the development of both atherosclerosis and acute coronary syndrome (ACS). The activation of platelets is the primary mechanism of thrombosis and ischemic events; thus, their inhibition represents an elective therapeutic approach for the treatment of atherothrombotic diseases. Several antiplatelet drugs with different characteristics have demonstrated their efficacy in reducing major adverse cardiovascular events.

25.1.1 Platelets, Atherothrombotic Diseases, and Antiplatelet Therapy

Platelet activation plays a key role in the development of both atherosclerosis and acute coronary syndrome (ACS) [1,2]. It is accepted knowledge that platelets adhere to the damaged walls of blood vessels at sites of endothelial cell activation, contributing to the development of chronic atherosclerotic plaques, and trigger the acute onset of arterial thrombosis in response to atherosclerotic plaque rupture [2].

Platelets are 2–3 μm diameter, disk-shaped cell fragments that derive from fragmentation of megakaryocytes and circulate in the bloodstream of humans for 7–10 days [3]. They play a crucial role in primary hemostasis, where their rapid response to vascular injury prevents excessive blood loss, but they are also a critical factor in the onset of pathological manifestations. In fact, an alteration of platelet number and/or phenotype determines thrombus formation, which is responsible for the clinical manifestations of atherothrombotic disease: ACS (unstable angina, or UA, non-ST-elevation myocardial infarction, or NSTEMI, and

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00025-6
© 2014 Elsevier Inc. All rights reserved.

ST elevation myocardial infarction, or STEMI), ischemic stroke, transient ischemic attack (TIA), and symptomatic peripheral arterial disease (PAD). These are the main causes of morbidity and mortality worldwide [3–5].

Platelet activation is a key process in both protective hemostasis and pathological thrombosis through the activation of multiple pathways by the binding of several agonists (e.g., thromboxane A2 (TxA2), adenosine diphosphate (ADP), and thrombin) to their receptors (Figure 25.1). Platelet adhesion to the subendothelium is mediated by direct interaction between the glycoprotein (GP)Ib/V/IX receptor complex on the platelet surface and the von Willebrand factor (vWF). Moreover, the interaction between exposed subendothelial collagen with platelet receptors GPVI and GPIa stimulates the release of platelet agonists ADP and TxA2 from the adherent platelets, as well as activation of the GPIIb/IIIa receptor with high affinity for fibrinogen, which mediates stable adhesion of platelets to the vessel wall, platelet–platelet cross-linking, and contact-dependent signaling within platelet aggregates (Figure 25.1) [1,3–5].

Release of ADP and TxA2 promotes the recruitment of circulating platelets into the growing stable hemostatic plug. Thrombin-mediated cleavage of fibrinogen into fibrin further contributes to the formation of hemostatic plugs.

ADP and TxA2 activate platelets by binding to specific receptors on the platelet-P2Y$_1$ and P2Y$_{12}$ receptors for ADP and prostaglandin G2 and H2 (PGG2 and PGH2) endoperoxide receptors (TPα and TPβ) for TxA2. These bindings result in reduced intracellular cyclic adenosine monophosphate levels and full activation of GPIIb/IIIa. ADP and TxA2 can also potentiate platelet activation induced by other ligands (Figure 25.1). Thrombin activates platelets primarily by binding the protease-activated receptor (PAR)-1 on the platelet surface, cleaving the receptor, and exposing a tethered ligand that binds and activates the receptor (Figure 25.1). Thrombin is the most potent platelet agonist, as it can stimulate platelet activation via the PAR-1 at very low concentrations that are several orders of magnitude lower than those required for the activation of the coagulation cascade. Human platelets also express a secondary receptor for thrombin, PAR-4, which requires higher concentrations of thrombin for activation [1,3–5].

Other factors, such as epinephrine, PGE2, serotonin, and several chemokines play a role in platelet activation, mainly potentiating platelet activation induced by other stimuli. On the other hand, ligand-stimulated platelet activation is inhibited by a number of endothelial-derived factors, nitric oxide, and prostacyclin, which in physiological conditions prevent uncontrolled platelet aggregation and cause an increase in the intracellular levels of cyclic nucleotides (cyclic guanosine monophosphate and cyclic adenosine monophosphate). In addition, the nucleoside adenosine—released as a result of cell damage or by endothelial ectonucleotidase

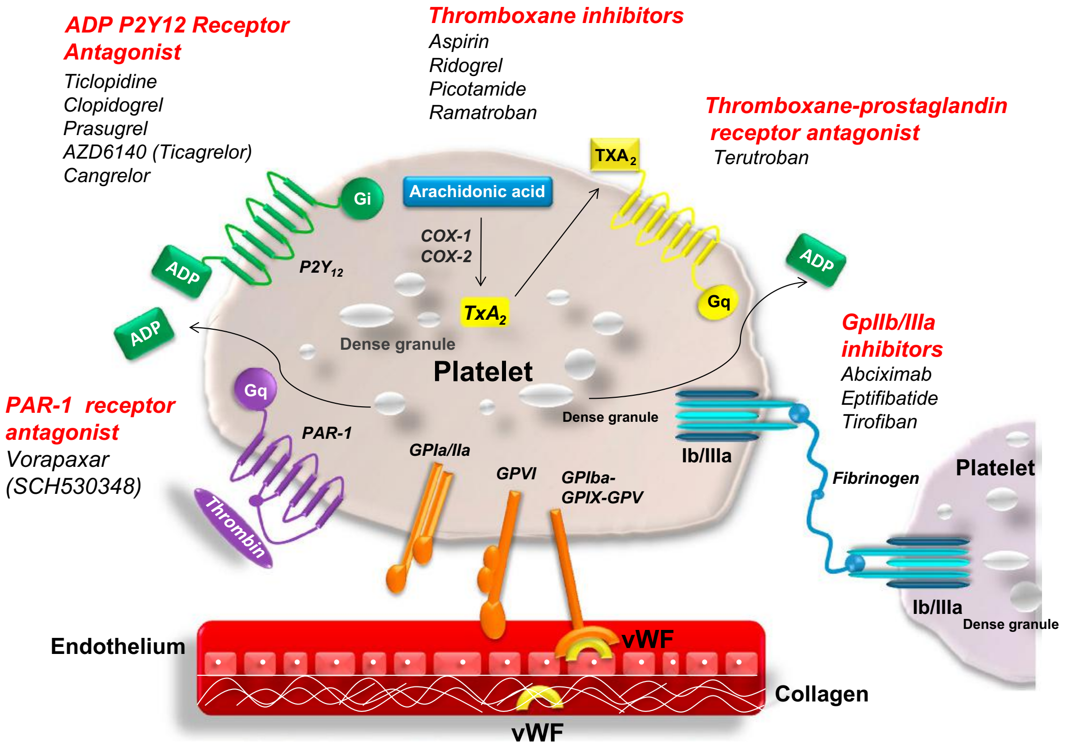

ADP P2Y12 Receptor Antagonist
Ticlopidine
Clopidogrel
Prasugrel
AZD6140 (Ticagrelor)
Cangrelor
Thromboxane inhibitors
Aspirin
Ridogrel
Picotamide
Ramatroban
Thromboxane-prostaglandin receptor antagonist
Terutroban
TXA_2
Gi
Arachidonic acid
COX-1
COX-2
$P2Y_{12}$
ADP
ADP
TxA_2
Gq
ADP
Dense granule
Platelet
Gq
PAR-1 receptor antagonist
Vorapaxar (SCH530348)
GpIIb/IIIa inhibitors
Abciximab
Eptifibatide
Tirofiban
PAR-1
GPIa/IIa
GPVI
GPIba-GPIX-GPV
Dense granule
Ib/IIIa
Fibrinogen
Platelet
Thrombin
Ib/IIIa
Dense granule
Endothelium
vWF
Collagen
vWF

FIGURE 25.1 **Pathways in platelet activation and aggregation, and mechanisms of action of classes of antiplatelet drugs.** ADP—adenosine diphosphate; COX—cyclooxygenase; GP—glycoprotein; PAR-1—protease-activated receptor-1; TXA2—thromboxane A2; vWF—von Willebrand factor.

CD39-mediated conversion of ADP—also inhibits platelet activation via activation of the Gs-coupled adenosine A2A receptor [1,3–5]. Therefore, the perpetuation phase of thrombus formation is mediated by the cell-to-cell contact-dependent mechanisms that lead to changes in platelet morphology, expression of procoagulant and proinflammatory molecules, and platelet aggregation [1,3–5].

Thrombus in acute atherothrombotic events (ACS, ischemic stroke, or symptomatic PAD) can be either partially or completely occlusive—the former composed primarily of platelet aggregates; and the latter of a platelet aggregate core and a superimposed fibrin-rich clot generated by the coagulation cascade [4].

Because the activation of multiple platelet pathways, in particular TxA2 and ADP platelet activation pathways, is the primary mechanism of thrombosis and ischemic events, their comprehensive inhibition is an attractive therapeutic approach for the treatment of atherothrombotic diseases (Figure 25.1). On the other hand, the potential clinical benefits of targeting various platelet activation pathways should be carefully weighed against the likelihood of increased bleeding, as both the TxA2 and ADP platelet activation pathways are also required for hemostasis.

In the last decades, oral antiplatelet agents that inhibit platelet activation by targeting cyclooxygenase 1 (COX-1) inhibition of TxA2 formation (acetyl salicylic acid, ASA, or aspirin) or by targeting the ADP-induced $P2Y_{12}$ receptor pathway (ADP $P2Y_{12}$ receptor inhibitors or thienopyridines, such as clopidogrel, ticlopidine, prasugrel, and ticagrelor) have been deeply studied and have been demonstrated to significantly reduce the incidence of ischemic events in patients with atherothrombotic diseases; also, they are widely used in clinical practice (Figure 25.1) [4,6]. The antiplatelet drugs panel also includes phosphodiesterase inhibitors (cilastozol), GPIIb/IIIa inhibitors (abciximab, eptifibatide, tirofiban), adenosine reuptake inhibitors (dipyridamole), and thromboxane inhibitors (thromboxane synthase inhibitors, thromboxane receptor antagonists) (Figure 25.1).

Antiplatelet drugs are widely used in primary and secondary prevention of thrombotic cardiovascular and cerebrovascular disease, according to current guidelines [2,4,6–8].

In particular, clopidogrel is indicated for prevention of vascular ischemic events in patients with symptomatic atherosclerosis, NSTEMI and STEMI, and it is used along with ASA for the prevention of thrombosis after percutaneous coronary intervention (PCI) with placement of an intracoronary stent. It is also used as an alternative antiplatelet drug for patients who are intolerant of ASA or with a history of gastric ulceration [2,4,6–8].

A large body of evidence exists showing that dual antiplatelet therapy (aspirin and clopidogrel) has a central role in reducing the risk of recurrent cardiovascular events in patients with coronary artery disease (CAD). Dual antiplatelet therapy is the standard of care for acute coronary syndrome (ACS) patients, who are managed either medically or by PCI [4,6]. In particular, clopidogrel treatment has delivered substantial benefits for patients undergoing PCI and stent implantation. Those receiving bare-metal and drug-eluting stents (BMS and DES) are recommended for clopidogrel-maintaining therapy for 1 month and at least 12 months, respectively. Retrospective studies have shown that the discontinuation of clopidogrel after stent implantation is associated with an increased risk of thrombotic events in patients with DES [4,6,9].

However, major adverse cardiovascular events (MACE)—namely, myocardial infarction (MI), cardiovascular death, stent thrombosis (ST), and stroke—can also occur in patients on clopidogrel therapy. Several retrospective and prospective studies have shown that, in patients who suffer MACE, the presence of on-clopidogrel high-platelet reactivity (HPR) might be demonstrated [10–15].

Consistent evidence has been accumulating in the literature on the prognostic role of HPR [16–19]. This has led to the identification of ADP-induced HPR as a marker of vascular risk in an international consensus document [20].

All these data suggest that variable response to clopidogrel is one of the main causes of thrombotic events in high-risk vascular patients. Considering the wide spectrum of alternative antiplatelet drugs, a deep knowledge of genetic and nongenetic determinants of clopidogrel response variability is crucial to the development of a personalized therapeutic strategy that targets individual patient characteristics.

25.1.2 Clopidogrel Metabolism

Different agonists induce platelet aggregation and fibrinogen binding to the subsequently activated GPIIb/IIIa receptor. In this process, ADP is a pivotal actor because it is released by damaged cells and activated platelets, thus enhancing the action of other platelet activators. ADP mediates platelet aggregation through its binding to two G protein-coupled receptors, $P2Y_1$ and $P2Y_{12}$, acting together to achieve complete aggregation. $P2Y_{12}$ also has an important role in amplifying responses to other agonists. In fact, ADP is released from platelet-dense granules regardless of the activating stimulus, be it thrombin, collagen, thromboxane A2, or any other agonist [1,3–5]. Because of the major role of $P2Y_{12}$ in arterial thrombosis and concomitant inflammatory response, the pharmacological targeting of this receptor has become an important pharmacological strategy in cardiovascular diseases. In fact, several antiplatelet drugs of the thienopyridine class have been developed that are capable of selectively binding to and inhibiting platelet $P2Y_{12}$ receptors (ticlopidine, clopidogrel, prasugrel, ticagrelor) (Figure 25.2).

Clopidogrel is a second-generation thienopyridine that has gradually replaced first-generation ticlopidine because

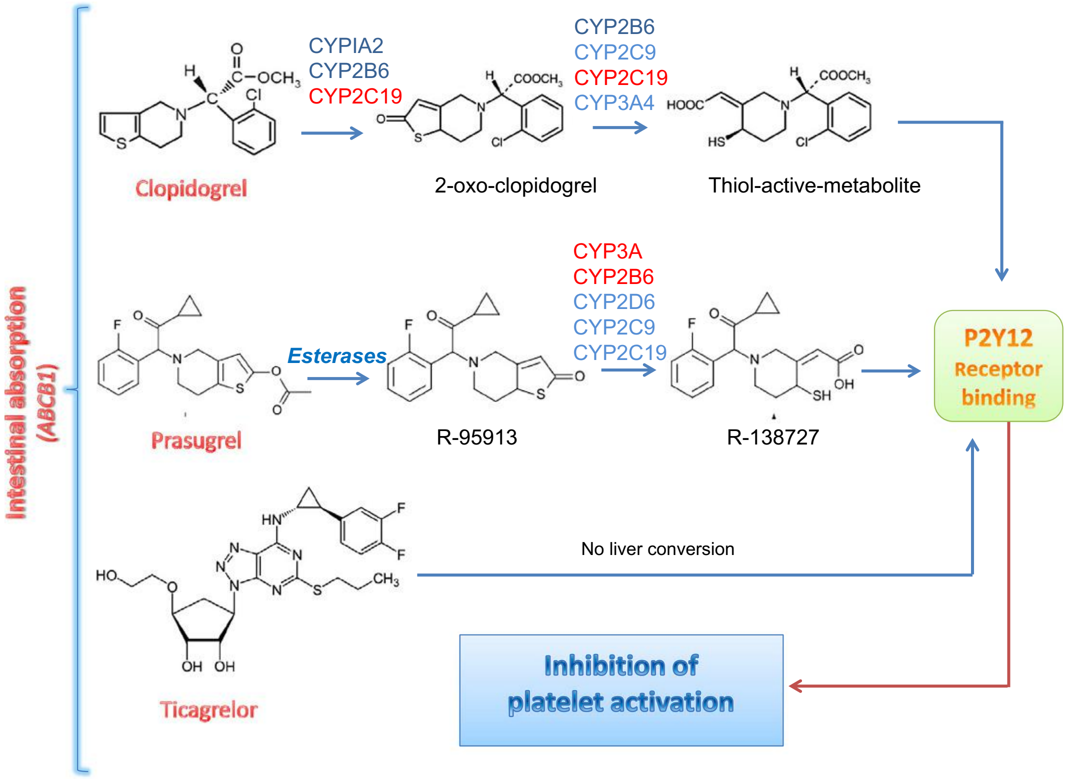


FIGURE 25.2 **Mechanism of action and metabolic pathways for clopidogrel, prasugrel, and ticagrelor.** Clopidogrel and prasugrel irreversibly bind the ADP P2Y$_{12}$ receptor. Ticagrelor reversibly binds the ADP P2Y$_{12}$ receptor.

of its freedom from serious toxicity, such as neutropenia and thrombotic thrombocytopenia purpura. The clinical efficacy of clopidogrel has been demonstrated both as an add-on to aspirin in the clinical setting, particularly for ACS patients undergoing PCI and stent implantation, and as a single antiplatelet therapy for secondary prevention [4,6].

Clopidogrel selectively inhibits the binding of ADP to the P2Y$_{12}$ subtype of the ADP receptor, and consequently inhibits the activation of the GPIIb/IIIa complex mediated by ADP. Therefore, it inhibits platelet aggregation. Because of the irreversible binding, platelets exposed to clopidogrel are affected for the rest of their life (about 7–10 days) and the recovery of platelet function occurs with a normal platelet turnover trend. Even platelet aggregation induced by agonists other than ADP is inhibited by blocking the amplification of platelet activation by released ADP. Platelet inhibition can be demonstrated 2 hours after a single dose of oral clopidogrel, but the onset of action is slow so that a loading dose of 300 or 600 mg is usually administered, followed by a daily maintenance dose of 75 mg [4,6,20].

Clopidogrel is a prodrug requiring hepatic metabolism (oxidation) by cytochrome P450 (CYP450) enzyme isoforms for its conversion to the active thiol metabolite SR26334 (Figures 25.2 and 25.3) [21–23]. It is absorbed in the duodenum, where the P-glycoprotein, coded by the *ABCB1* gene, member 1 of the ATP binding cassette subfamily B (MDR/TAP), is involved in the intestinal transport (Figure 25.3). Only about 15% of the clopidogrel prodrug undergoes metabolism by the CYP450 enzyme system in the liver; the remainder (85%) is hydrolyzed by esterases to an inactive carboxylic acid derivative (Figure 25.3) [23].

Several CYP450 enzyme isoforms, with different contributions, are responsible for the oxidation of the thiophene ring of clopidogrel (Figures 25.2 and 25.3). The first oxidation step to 2-oxo-clopidogrel is mediated by the enzymes CYP1A2 (35.8%), CYP2B6 (19.4%), and CYP2C19 (44.9%), and the intermediate metabolite undergoes further oxidation by the enzymes CYP2B6 (32.9%), CYP2C9 (6.76%), CYP2C19 (20.6%), and CYP3A4 (39.8%). The final product has the opened thiophene ring, resulting in formation of both a carboxyl and a thiol group (Figures 25.2 and 25.3) [9,21,24]. The clopidogrel active metabolite forms a disulfide bridge with the two extracellular cysteine residues located on the ADP P2Y$_{12}$ receptor, and causes an irreversible blockade of ADP binding for the platelet's lifespan [25]. Data on clopidogrel metabolism emphasize that CYP2C19 contributes substantially to

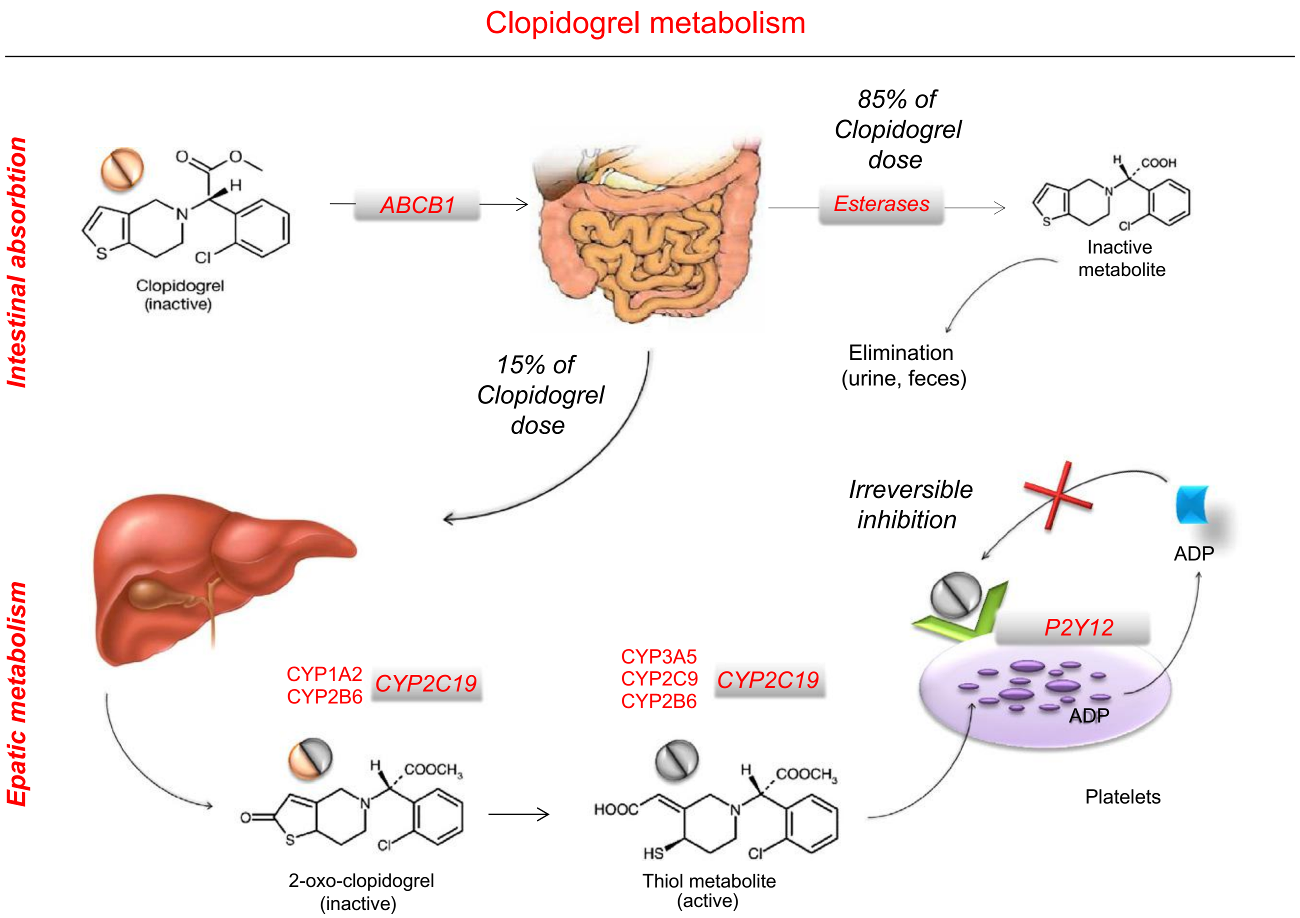

Clopidogrel metabolism
Intestinal absorbtion
Clopidogrel (inactive)
Cl
S
N
H
O
ABCB1
85% of Clopidogrel dose
Esterases
H
COOH
Inactive metabolite
Elimination (urine, feces)
15% of Clopidogrel dose
Irreversible inhibition
ADP
P2Y12
ADP
Platelets
Epatic metabolism
CYP1A2
CYP2B6
CYP2C19
CYP3A5
CYP2C9
CYP2B6
CYP2C19
COOCH3
2-oxo-clopidogrel (inactive)
HOOC
HS
Thiol metabolite (active)

FIGURE 25.3 **Clopidogrel metabolism.** Clopidogrel is absorbed intestinally. In this step, P-glycoprotein, encoded by the *ABCB1* gene, is actively involved. Clopidogrel is mostly hydrolyzed by esterases to an inactive carboxylic acid derivative that accounts for 85% of clopidogrel-related circulating compounds. Several cytochrome P450 (CYP450) enzyme isoforms are responsible for oxidation of the thiophene ring of clopidogrel to 2-oxo-clopidogrel, which is further oxidized by other CYP450 isoforms to cause the opening of the thiophene ring and formation of the carboxyl and thiol groups. The thiol group irreversibly binds to the ADP $P2Y_{12}$ receptor expressed on platelet surfaces and causes irreversible blockade of ADP binding. *ABCB1*—ATP-binding cassette subfamily B (MDR/TAP, member 1 gene coding for the P-glycoprotein involved in clopidogrel intestinal absorption); CYP1A2—cytochrome P450 (family 1, subfamily A, polypeptide 2); CYP2B6—cytochrome P450 (family 2, subfamily B, polypeptide 6); CYP2C19—cytochrome P450 (family 2, subfamily C, polypeptide 19); CYP2C9—cytochrome P450 (family 2, subfamily C, polypeptide 9); CYP3A4—cytochrome P450, family 3, subfamily A, polypeptide 4); P2Y12—purinergic receptor P2Y (G-protein coupled, 12).

both of the oxidative steps required in the formation of the clopidogrel active metabolite (Figures 25.2 and 25.3). Understanding this phenomenon clarifies the significant effect of genetic polymorphisms of the *CYP2C19* gene, which alter the function of this important CYP450 isoform. Indeed, among the different CYP450 isoenzymes involved in clopidogrel metabolism, single-nucleotide polymorphisms (SNPs), such as the *CYP2C19*2* loss-of-function polymorphism, have been associated with the CYP2C19 enzyme's inactivation and with impaired clopidogrel metabolism.

Moreover, the CYP2C19 enzyme is responsible for the metabolism and clearance of numerous drug substrates including proton pump inhibitors (PPIs), diazepam, phenytoin, and cyclophosphamide (http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm). Therefore, possible drug–drug interactions should be considered in complex patients taking multiple medications.

25.1.3 Alternative Thienopyridine-Class Chemical Compounds

Other chemical compounds of the ADP receptor inhibitors thienopyridine class, prasugrel and ticagrelor, are alternatives to clopidogrel for antiplatelet therapy in the clinical setting of high-risk ACS patients undergoing PCI and stent implantation (Figure 25.2) [6].

25.1.3.1 Prasugrel

Prasugrel, a third-generation thienopyridine, is administered in ACS patients managed with PCI, and requires a loading dose to achieve a rapid, maximal antiplatelet effect. It reduces the aggregation of platelets by binding to the $P2Y_{12}$ ADP receptor. Prasugrel is a prodrug and is rapidly metabolized into a pharmacologically active form. It is hydrolyzed by carboxyesterases to a thiolactone, R-95913, which is subsequently metabolized to the prasugrel-active metabolite, R-138727. Conversion of R-95913 to R-138727 is catalyzed by several CYP450 enzymes, with the greatest contributions from CYP3A and CYP2B6, lesser contributions from CYP2C9 and CYP2C19, and few from CYP2D6. The active metabolite of prasugrel irreversibly binds to the ADP $P2Y_{12}$ receptor (Figure 25.2) [4,6,26,27].

Prasugrel is associated with greater active metabolite generation than is clopidogrel, and its maximum plasma-active metabolite concentration is reached in approximately 30 min [26]. Compared to standard and higher doses of clopidogrel, prasugrel inhibits ADP-induced platelet aggregation more rapidly, more consistently, and to a greater extent in healthy volunteers and in patients with CAD, including those undergoing PCI. Prasugrel is characterized also by less response variability than clopidogrel [4,6,28]. In the TRITON-TIMI 38 study (Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition-Thrombolysis in Myocardial Infarction)—randomized, double-blind clinical trial involving 13,608 patients with ACS undergoing PCI—the prevalence of cardiovascular death, nonfatal myocardial infarction, or stroke was lower with prasugrel than with clopidogrel. However, rates of bleeding were higher in the prasugrel group [29].

25.1.3.2 Ticagrelor

Ticagrelor (AZD6140) is the first generation of a new class of reversible $P2Y_{12}$ receptor inhibitors called cyclopentyl-triazolopyrimidines. Like the thienopyridines, prasugrel, clopidogrel, and ticlopidine, ticagrelor blocks the $P2Y_{12}$ ADP receptor, but unlike the other thienopyridines, the block is reversible (Figure 25.2) [6,27,30]. Furthermore, the molecule does not require activation by the liver, which reduces the risk of drug interactions, resulting in rapid onset of action with less interindividual variation. It also inhibits nonplatelet $P2Y_{12}$ receptors present on vascular smooth muscles, where they cause vasoconstriction and increase myocardial perfusion. Ticagrelor binds to an area distinct from the ADP binding site and inhibits ADP-induced receptor activation in a noncompetitive manner. It leaves the receptor intact on dissociation, so inhibition of platelet aggregation is reversible (Figure 25.2) [6,30]. Ticagrelor was evaluated against clopidogrel in patients with ACS in the PLATO (A Study of Platelet Inhibition and Patient Outcomes) trial [31]. It was associated with significant reduction in cardiovascular death, MI, and stroke, with no difference in the overall incidence of major bleeding (defined by PLATO researchers), but with an increase in major bleeding related to non-coronary-artery bypass graft [31].

Moreover, other antiaggregant drugs, such as a PDE3 phosphodiesterase inhibitor (cilostazol), and an additional short-acting competitive $P2Y_{12}$ platelet receptor antagonist (cangrelor) are available and might be considered in the management of high-risk vascular patients.

25.2 CLOPIDOGREL RESPONSE VARIATION

Platelet inhibition by clopidogrel is highly variable, and patients with reduced platelet inhibition have an increased risk for MACE [20]. In fact, a growing number of studies have linked high platelet reactivity despite clopidogrel treatment to an increased risk of MACE, in particular ST [16,17,19,32]. Sofi and co-workers conducted a meta-analysis of the clinical consequences of clopidogrel nonresponsiveness among more than 4500 CAD patients who had undergone a PCI. They followed these patients for a period ranging from 2 weeks to 1 year after a PCI and found poor responders beat increased risk of cardiovascular clinical recurrences as opposed to those with a good response (OR 5.67 (95% CI 2.97–10.84), $p<0.00001$) [16].

Even if a single treatment strategy directed against a specific receptor cannot be expected to overcome all thrombotic events, and clinical treatment failure (e.g., an ischemic event) during clopidogrel treatment is not considered synonymous with clopidogrel resistance [20], an optimal level of platelet inhibition is a sine qua non condition of cardiovascular event prevention without increased bleeding risk. At present, despite numerous papers evaluating this issue, there is no platelet function cut-off that is widely accepted in identifying patients at high risk for ischemic events in different clinical settings [20,32].

25.2.1 High on-Clopidogrel Platelet Reactivity

Platelet function can be evaluated by several laboratory methods. These include the following:

- Light transmission aggregometry on platelet-rich plasma (PRP LTA)
- VerifyNow (Accumetrics, San Diego, CA; formerly known as Ultegra Rapid Platelet Function Assay)
- Platelet Function Analyzer 100 (PFA-100, Dade-Behring, Marburg, Germany)
- Flow cytometry
- Whole-blood aggregometry by multiple electrode platelet aggregometry (Multiplate Analyzer, Dynabyte, Munich, Germany)

Among methods for evaluating on-clopidogrel platelet reactivity, PRP LTA induced by ADP [33,34] is the gold standard. In brief, aggregometry evaluates the complex response of the platelet to ADP stimulation through the function of both $P2Y_1$ and $P2Y_{12}$ receptors. There have been studies on the maximal amplitude of measured platelet aggregation in response to 5, 10, or 20 μM ADP [20]. Despite the wide use of aggregometry, because of the many manual phases involved in the procedure, the LTA is characterized by poor standardization between institutions; thus, at present it may not be the ideal test to monitor the effects of antiplatelet therapy outside of specialized laboratories. Flow cytometry using the vasodilator-stimulator phosphoprotein (VASP) assay detects the availability of the $P2Y_{12}$ receptor and determines the effectiveness of clopidogrel. Among point-of-care tests to be used at the patient bedside or in less specialized laboratories are the following:

- PFA-100, which uses cartridges with epinephrine or ADP to mimic a high-shear state to evaluate platelet function
- The Verify Now assay system, a turbidimetry-based optical device widely used in clinical trials for testing on-clopidogrel platelet aggregation inhibition.

Paniccia and co-workers compared laboratory methods—VerifyNow $P2Y_{12}$, PFA-100, VASP, multiple electrode platelet aggregometry on whole blood, and PRP LTA—and found a significant correlation between the classical PRP LTA and the other platelet function testing platforms [33–36].

POPular was a large study that compared different platelet function methods in relation to their capacity for predicting cardiovascular events. In this study, LTA and VerifyNow were demonstrated to predict cardiovascular events, even with a low positive predictive value [37].

Several studies have been published confirming the association between on-clopidogrel HPR (regardless of the test used for its definition) and enhanced risk of adverse clinical events [16,17,19,32]. In several studies, such as RECLOSE (Low Responsiveness to Clopidogrel and Sirolimus- or Paclitaxel-Eluting Stent Thrombosis) [11,12] and RECLOSE2-ACS [18], on-clopidogrel HPR was defined at 10 μM ADP-induced maximal LTA ≥70%.

Much attention has been focused on clinical validation of point-of-care tests' determination of the cut-off for definition of HPR. For example:

- Price and co-workers conducted a study of 380 patients undergoing PCI with stent implantation. They found that the optimal cut-off for the definition of on-clopidogrel HPR associated with adverse clinical events, at a 6-month follow-up, was ≥235 PRU ($P2Y_{12}$ reaction unit) [13], as measured by the VerifyNow$P2Y_{12}$ assay.
- In the ARMYDA-PRO study (Antiplatelet Therapy for Reduction of Myocardial Damage during Angioplasty—Platelet Reactivity Predicts Outcome), and in the Marcucci and co-workers study, both using VerifyNow, the cut-off was ≥240 PRU associated with: (1) 30-day MACE on 160 patients undergoing PCI cardiovascular death; and (2) nonfatal MI at 12-month follow-up on 683 ACS patients, respectively [14,38].
- In the GRAVITAS trial (formally, Gauging Responsiveness with A VerifyNow Assay—Impact on Thrombosis and Safety), which followed 2214 patients, Price and co-workers defined on-clopidogrel HPR at ≥230 PRU [39].

25.2.2 Nonpharmacogenetic Mechanisms of High on-Clopidogrel Platelet Reactivity

Failure of clopidogrel response is due to multiple chronic or transient mechanisms. Chronic mechanisms include inadequate drug compliance, drug–drug interactions, age, diabetes mellitus, elevated body mass index (BMI), and reduced left ventricular ejection fraction (LVEF) (Figure 25.4). Transient mechanisms include inflammation, accelerated platelet turnover, reticulated platelets (RPs), erythrocyte deformability, and the activity of a disintegrin and metalloproteinase with thrombosporin motif-13(ADAMTS13) (Figure 25.4) [23]. All of these mechanisms could interact with genetic predisposition in determining inadequate response to clopidogrel and consequently increased risk of MACE, including ST, in high-risk vascular patients.

25.2.2.1 Compliance

Clopidogrel noncompliance may represent a critical issue. In a cohort of patients after stroke, discontinuation of clopidogrel therapy was observed in ~15% (within 30 days) of patients with CAD [40], in 18.4% (at 3 months), and in 38.4% (at 1 year) [41]. Whereas noncompliance represents a minor issue in short-term studies, it may have a significant impact on longer follow-up studies. In CAD and post-ischemic stroke patients, therapy with clopidogrel was associated with double-digit underestimated risks for noncompliance, especially in stroke survivors. This supports the hypothesis that lack of inhibition of platelet aggregation and clopidogrel resistance are attributed to hidden noncompliance [42]. Much evidence tells us that discontinuation of clopidogrel is associated with a significantly higher risk of ischemic events—first of all, ST [43,44]. Recently, a study assessing the risk associated with temporary (median: 7 days) discontinuation of double antiplatelet therapy, during the first year after DES implantation in 1622 consecutive patients, showed that 10.6% interrupted at least one antiplatelet drug during the first year after DES implantation (only one during the first month). Discontinuation was followed by ACS in 4.1% of patients—a rate of major cardiac

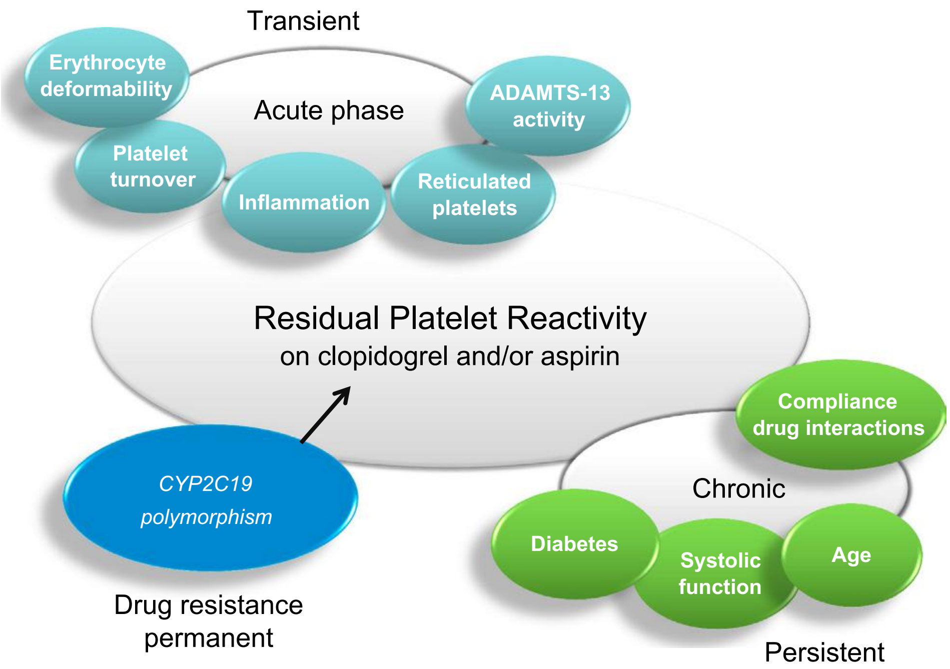


FIGURE 25.4 Mechanisms of clopidogrel response variability.

events similar to that in patients without antiplatelet therapy discontinuation (5.5%). Thus, the researchers observed that antiplatelet therapy discontinuation for a few days was not independently associated with 1-year major cardiac events (hazard ratio: 1.32 (95% CI 0.56–3.12)) [45].

25.2.2.2 Clopidogrel Interaction with Other Drugs

Other drugs taken at the same time may interfere with clopidogrel inhibition of platelet reactivity. This phenomenon may derive from high doses of calcium-channel blockers and angiotensin-converting enzyme inhibitors [9]. The lipophilic statin atorvastatin is frequently studied in clopidogrel trials because it has a high affinity for the CYP3A4 isoenzyme. Studies evaluating clopidogrel–atorvastatin interaction have not given consistent results. Some reported that atorvastatin promoted clopidogrel resistance. On the other hand, pravastatin or rosuvastatin, characterized by a nonCYP3A4 metabolism, did not seem to significantly affect clopidogrel platelet inhibition. In fact, in PCI-treated patients with HPR during co-administration of clopidogrel and atorvastatin, switching to a nonCYP3A4-metabolized statin can significantly decrease platelet reactivity and the prevalence of HRP. This effect appears similar irrespective of the type of nonCYP3A4-metabolized statin [46].

More recent studies, however, did not confirm this interaction. Post hoc analysis of several important trials—for example, CREDO (Clopidogrel for the Reduction of Events During Observation), PRONTO (Plavix Reduction of New Thrombus Occurrence), and MITRAPLUS (Maximal Individual Therapy in Acute Myocardial Infarction Plus—have not revealed a significant interaction between statins (i.e., atorvastatin, fluvastatin, lovastatin, pravastatin, simvastatin) and clopidogrel in terms of platelet inhibition [9].

The results of a recent prospective pharmacodynamic randomized study known as ACHIDO (Atorvastatin and Clopidogrel High Dose) that was conducted on stable patients with residual high platelet activity showed that treatment with high-dose atorvastatin (80 mg) associated with double-dose clopidogrel (150 mg) reduces platelet reactivity significantly more than does double-dose clopidogrel alone in statin-naïve patients with stable CAD and HPR before PCI in both diabetes mellitus and nondiabetes mellitus [47,48]. This datum suggests a possible role for statins in decreasing platelet reactivity through its anti-inflammatory effect independently of the hepatic pathway of metabolization. In particular, the antiplatelet effects of high-dose atorvastatin occur early and persist over time; they are additive to those of clopidogrel and do not correlate with the degree of lipid reduction observed during the study period [48].

There is a growing body of evidence on the interaction between the proton pump inhibitors (PPIs) and clopidogrel in influencing platelet reactivity and adverse outcomes [49–51]. It is estimated that ~15% of an absorbed clopidogrel dose is converted to an active thiol metabolite, mainly by the hepatic CYP450 isoenzymes [22]. One of these isoenzymes, the CYP2C19 isoform, is involved in the metabolism of many PPIs, which are also, to varying degrees, CYP2C19 isoenzyme inhibitors. The ability of PPIs to inhibit CYP2C19 activity reduces the active metabolite of clopidogrel generation and causes a diminished platelet response to clopidogrel. In addition, PPIs may influence antiplatelet clopidogrel response through the variability of intestinal clopidogrel absorption [9]. PPIs are substrates and inhibitors of the intestinal efflux transporter P-glycoprotein, a key factor in intestinal absorption of clopidogrel. Alterations in P-glycoprotein activity can potentially affect plasma concentrations of clopidogrel, and hence the concentration of active metabolite [9].

Gilard and co-workers documented for the first time that the contemporary administration of clopidogrel and omeprazole is associated with reduced platelet inhibition related to the $P2Y_{12}$ receptor in CAD patients. Those undergoing stent implantation on dual antiplatelet treatment in the OCLA (Omeprazole Clopidogrel Aspirin) trial were randomized to receive omeprazole 20 g/day or placebo for 7 days. This treatment significantly decreased the effect of clopidogrel on platelet reactivity as tested by platelet vasodilator-stimulated phosphoprotein (VASP) phosphorylation assay [52]. The influence of pantoprazole and esomeprazole on clopidogrel antiplatelet response was assessed in different studies, which reported that the intake of these two drugs was not able to impair platelet inhibition by clopidogrel [9]. The effect of pantoprazole and omeprazole on inhibition by clopidogrel was evaluated in 104 non-ST-segment elevation (NSTE) ACS patients. This study found, after 1 month of treatment, a significantly reduced platelet inhibition by omeprazole but not by pantoprazole [53].

Contrasting data about the capability of PPI administration to affect the risk of clinical events are also available [49–51]. In 13,636 patients with ACS treated with clopidogrel, the use of PPIs was associated with an increased risk of acute MI [54]. However, the use of pantoprazole did not significantly affect the risk of acute MI, whereas the other PPIs were associated with an increased risk of about 40%. In a retrospective study of 8205 patients with ACS, the use of clopidogrel plus PPIs (at discharge, during follow-up, or both) was associated with an increased risk of death or rehospitalization for ACS compared with the use of clopidogrel without PPIs. By analyzing the different PPIs in relation to adverse outcomes, the earlier mentioned study failed to demonstrate a significant difference between omeprazole and raboprazole [55].

The PRINCIPLE-TIMI 44 study (formally, Prasugrel in Comparison to Clopidogrel for Inhibition of Platelet Activation and Aggregation—Thrombolysis in Myocardial Infarction 44) and the TRITON-TIMI 38 study (formally, the Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition—Thrombolysis in Myocardial Infarction 38) retrospectively analyzed data in order to verify the possible interactions between the use of PPIs, platelet function, and adverse cardiovascular events in relation to both clopidogrel and prasugrel. In both studies, the use of PPIs was not significantly associated with an increased risk of ischemic events in patients treated with clopidogrel and prasugrel; this was confirmed in the subgroup analysis with the single PPI (omeprazole, pantoprazole, esomeprazole, and lansoprazole) in both sets of patients [56]. Recently, Parri and co-workers published the first documentation in a randomized trial showing that, after correction for the bias of *CYP2C19*2* polymorphism, pantoprazole increases the ADP-maximal aggregation in patients treated with dual antiplatelet therapy [57]. In addition, the PLATO study retrospectively evaluated the possible influence of PPIs on the clopidogrel and ticagrelor arm. Surprisingly, they found a significantly higher risk of CV events associated with both drugs. Thus, the association of PPIs and adverse events may be due to confounding factors—namely, that patients administered PPIs have a higher morbidity—and so PPIs might be more a marker than a cause of higher rates of cardiovascular events [58].

25.2.2.3 Age

The available evidence demonstrates that advanced age is associated with a significantly high percentage of on-clopidogrel HPR. One of the causes may be the fact that the conversion of clopidogrel to its active metabolite is impaired in older patients [9]. In a prospective study that evaluated 191 patients who underwent PCI, it was shown that ADP-inducible platelet reactivity increased linearly with age after adjustment for cardiovascular risk factors, type of intervention, medication, C-reactive protein, and renal function [59]. At variance, the study by Price and co-workers evaluating various confounding factors in patients on continuous clopidogrel therapy demonstrated no effect on ADP-inducible platelet reactivity due to age (≥75 years) [60]. This discrepancy may be attributable to the duration of clopidogrel therapy. Most patients in the study by Gremmel and co-workers received a loading dose of clopidogrel prior to angioplasty, while the Price study assessed platelet aggregation mainly in patients on continuous clopidogrel therapy [59,60]. Therefore, the attenuation of clopidogrel-mediated platelet inhibition at an advanced age may be a phenomenon of the initial phase of antiplatelet therapy.

The SENIOR-Platelet study compared platelet function in 1331 elderly patients (>75) on thienopyridine (clopidogrel 75/150 mg/prasugrel 10 mg), versus that in younger patients (<75 years). It found that elderly patients present an impaired response to clopidogrel with a high rate of HPR. Clopidogrel 150 mg or prasugrel 10 mg blunted but did not eliminate the difference in response observed in old and young patients [61].

25.2.2.4 Diabetes Mellitus

Diabetes mellitus is a clinical condition characterized by enhanced platelet reactivity that exposes diabetic patients to an increased risk of atherothrombotic events. In diabetics, platelets are more frequently activated by subthreshold stimuli and consumed more rapidly, thus contributing to accelerated thrombopoiesis compared to platelets of nondiabetic individuals [62]. It has been documented that on-clopidogrel HPR is more prevalent in diabetic than in nondiabetic patients. In particular, type 2 diabetic patients undergoing elective PCI have approximately a four-fold increase in nonresponse at 24 hours following a standard loading dose. Moreover, platelet reactivity is persistently elevated in diabetic patients even in the maintenance phase of clopidogrel and is highest among those requiring insulin therapy [9,62].

25.2.2.5 Inflammation

Recent studies have documented a significant and independent association between platelet hyper-reactivity and both inflammatory markers (such as erythrocyte sedimentation rate and leukocyte number) and unbalanced proinflammatory/anti-inflammatory cytokine levels [63]. In ACS patients on dual antiplatelet treatment, during the acute phase of the disease, proinflammatory cytokine levels not sufficiently compensated by anti-inflammatory cytokines are able to modulate platelet hyper-reactivity [64]. Interestingly, Antonino and co-workers demonstrated that patients who underwent elective PCI treated with long-term clopidogrel and aspirin dual therapy had less inflammation, as indicated by specific biomarkers (C-reactive protein and other proinflammatory cytokines), compared with clopidogrel-naïve patients, suggesting that long-term clopidogrel therapy is associated with an anti-inflammatory effect [65].

25.2.2.6 Platelet Turnover and Reticulated Platelets

An elevated turnover of the circulating platelet pool yields a larger population of young platelets that are more reactive than older platelets. It has been shown that younger platelets have increased mean volume and a greater number of dense granules than do older circulating platelets [66]. In both healthy subjects and patients with stable CAD, it has been demonstrated that an increased platelet turnover, as reported by the proportion of circulating RPs, is associated with a significantly increased aggregation and activation of platelets even after dual antiplatelet therapy [9,66]. In 372 ACS patients on dual antiplatelet therapy, the number with suboptimal or “low” response to aspirin, clopidogrel, or both was clustered among patients with a high proportion of circulating RPs [67]. Moreover, RPs have been found to be independent predictors of cardiovascular death and may be useful per se in improving risk stratification in ACS patients [68].

25.2.2.7 The von Willebrand Factor and ADAMTS-13

The adhesion and activation of platelets at sites of vascular injury is mediated by the von Willebrand factor (vWF), which is released from storage sites or secreted by endothelial cells as ultralarge vWF (ULvWF) and rapidly cleaved through intravascular proteolysis by ADAMTS-13. In the acute phase of CAD and after the vascular injury induced by a PCI, endothelial cells are activated and high amounts of ULvWF are released. ULvWF forms are able to bind the GPIb–IX complex to aggregate platelets. In physiological conditions, the large multimeric forms are rapidly cleaved by ADAMTS-13 on the surface of the endothelium, resulting in the normal multimeric vWF pattern. When ADAMTS-13 activity is reduced, the ULvWFs accumulate and can induce platelet aggregation [9]. On-clopidogrel HPR is associated with lower ADAMTS-13 activity accompanied by elevated vWF levels [69], suggesting a modulatory effect of the interplay between ADAMTS-13 and vWF on antiplatelet therapy response in patients with marked endothelial activation, such as those with ACS who have undergone PCI.

25.2.2.8 Red Blood Cells and Leukocytes

Red blood cells and leukocytes are significantly implicated into the pathogenesis of arterial thrombosis. Platelet–leukocyte interactions play an important role in both adhesion and activation of platelets, but red blood cells are also involved in the interactions between platelets and the vessel wall. In patients with ACS, the inhibition of platelets induced by antiplatelet therapy is modulated not only by direct action on platelets, but also by erythrocyte deformability and white blood cell count [70]. This suggests that all blood components, including red blood cells, should be taken into account.

Some hemorheologic variables are also associated with on-treatment HPR in patients with ACS on dual antiplatelet therapy. In particular, decreased erythrocyte deformability is significantly associated with increased platelet aggregation induced by ADP [71]. A possible explanation for this association is that erythrocytes with decreased deformability tend to release a larger amount of ADP into the circulation compared to those without decreased deformability; the ADP released can compete against clopidogrel for purinergic receptors present on the platelet surface [72].

25.3 GENETIC DISSECTION OF CLOPIDOGREL RESPONSE

25.3.1 Genetic Determinants of Clopidogrel Response Variability

Among mechanisms responsible for the on-clopidogrel HPR, the individual susceptibility conferred by genetic variants in different genes coding platelet components (glycoprotein (Gp)Ia (*ITGA2*), GpIIIa (*ITGB3*), GpIb alpha (*GP1BA*), GpVI (*GP6*), ADP receptor $P2Y_{12}$ (*P2Y12*) and $P2Y_1$ (*P2Y1*), P-selectin (*SELP*), cyclooxygenase-1 (*COX1*), and cyclooxygenase-2 (*COX2*), endothelial nitric oxide synthase (*eNOS*)), or genes coding for several CYP450 enzyme isoforms involved in clopidogrel metabolism, has been hypothesized, studied and in some cases consistently demonstrated [73–76].

In particular, several polymorphisms in genes coding different isoforms of CYP450 (*CYP3A4*, *CYP3A5*, *CYP1A2*, *CYP2B6*, *CYP2C9*, *CYP2C19*) or the glycoprotein P involved in clopidogrel intestinal absorption (*ABCB1*) have been studied as determinants of on-clopidogrel HPR. For the majority of the investigated genetic variants results are negative or inconsistent.

On the contrary, consistent data are available concerning loss- and gain-of-function genetic variants in the gene *CYP2C19* on chromosome 10, and in particular the CYP2C19*2 polymorphisms. Several *CYP2C19* variant alleles have been identified and exhibit functional impairment which phenotypically categorized subjects as poor metabolizers [76,77].

The power of the evidence on the genetic contribution of genetic variants in the gene coding for one of the enzyme isoforms of CYP450, *CYP2C19*, were so consistent that in March 2010, the U.S. FDA (Food and Drug Administration) added a boxed warning to the clopidogrel label to alert clinicians to the impact of the CYP2C19 genetic profile on the drug's efficiency, thereby leaving the decision to use genetic testing to the discretion of physicians (see for details http://www.fda.gov/NewsEvents/ Newsroom/ PressAnnouncements/ ucm204253.htm and http://www.fda.gov/Drugs/DrugSafety/ PostmarketDrugSafetyInformation forPatientsandProviders/ucm203888.htm).

The *CYP2C19*2* loss-of-function polymorphism (rs4244285) is a G681A nucleotide substitution in the junction of intron 4 and exon 5, which introduces a splicing defect, resulting in a truncated, nonfunctional protein, responsible for the poor metabolizer phenotype [76,78,79]. The *CYP2C19*2* variant allele has different minor allele frequencies in Caucasian (15%), Asian (30%) and African-American (17%) populations [76,78–80].

The *CYP2C19*3* polymorphism (rs4986893) is a G636A nucleotide substitution in exon 4 that creates a premature stop codon at amino acid position 212 resulting in a truncated, metabolically inactive protein and poor metabolizer phenotype [76,78,79]. The *CYP2C19*3* polymorphism is less frequent in Asian (5%) and rare in both Caucasian (0.04%) and African-American (0.4%) populations [76,78–80]. Other rare variant alleles contribute to determine the poor metabolizer phenotype, such as *CYP2C19*4* (rs28399504, A-to-G substitution in the initiation codon, resulting in a Met1Val amino acidic substitution) and *CYP2C19*5* (rs56337013, C1297T nucleotide substitution in exon 9, resulting in an Arg433Trp amino acidic substitution) [76]. A further allelic variant, *CYP2C19*17* (rs12248560) in the *CYP2C19* gene has been discovered which results in an increased enzyme function due to a nucleotide substitution (-806C/T) in the 5′-flanking region of the gene that causes increased transcription of *CYP2C19* [76,78,79].

In 2010, Sibbing and co-workers conducted a study of 1524 patients undergoing PCI after pretreatment with 600 mg clopidogrel with a 30 day follow-up. They showed that *CYP2C19*17* carrier status is significantly associated with enhanced response to clopidogrel, as evaluated by whole-blood ADP-induced (6.4 μM) platelet aggregation (β coefficient = −32.3, SE 9.5, p = 0.0006 in the multivariable linear regression model), and an increased risk of bleeding (OR = 1.85, 95% CI 1.19−2.86, p = 0.006 in the multivariable logistic regression model) [81]. No significant influence of the *CYP2C19*17* polymorphism on ST was observed in this study [81]. However, the association was confirmed by further evidence, suggesting the need for *CYP2C19*17* in the diagnostic panel of clopidogrel pharmacogenetic variants [76–79]. Table 25.1 lists the *CYP2C19* genetic variants with known enzymatic activity alteration, and their characteristics [81–92].

25.3.2 Association of *CYP2C19*2* with High on-Clopidogrel Platelet Reactivity in Healthy Subjects

Several studies have evaluated the impact of the *CYP2C19*2* polymorphism on clopidogrel pharmacokinetics and pharmacodynamics in healthy volunteers (Table 25.2). Hulot and colleagues showed that the *CYP2C19* loss-of-function allele is associated with a marked decrease in platelet responsiveness to clopidogrel in 28 young healthy male subjects treated for 7 days with clopidogrel 75 mg/day (Table 25.2) [93]. Fontana and colleagues replicated the evidence of this association in 94 healthy subjects by assessing ADP-induced platelet aggregation after one week of clopidogrel administration (300 mg the first day; 75 mg/day maintenance dose). They too showed that *CYP2C19*2* explained 10% of the variability in clopidogrel responsiveness when the analysis was adjusted for age, platelet count, hematocrit, collagen lag time, fibrinogen, and vWF levels (Table 25.2) [94]. Brandt and co-workers genotyped healthy subjects for genetic variants in several genes involved in clopidogrel metabolism (*CYP1A2*, *CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP3A4*, and *CYP3A5*). These subjects were participating in studies evaluating pharmacokinetic and pharmacodynamic response to prasugrel (60 mg, n = 71) or clopidogrel (300 mg, n = 74) [95]. In subjects treated with clopidogrel, the presence of the *CYP2C19*2* polymorphism was significantly associated with lower exposure to the clopidogrel active metabolite, lower inhibition of platelet aggregation

TABLE 25.1 CYP2C19 Genetic Variants with Known Enzymatic Activity Alteration

CYP2C19 Allele	Gene Region	dbSNP	Nucleotide Substitution	Protein	Minor Allele Frequency	Enzymatic Activity	Study
*1			Wild-type	—		Normal	Richardson et al. [82]
*2	Exon 5	rs4244285	c.681G>A	Splicing defect	Europeans: 0.155 Asians: 0.256 Africans: 0.144	None	de Morais et al. [83]
*3	Exon 4	rs4986893	c.636G>A	p.W212X	Europeans and Africans: rare Asian: 0.005	None	Fukushima-Uesaka et al. [84]
*4	Exon 1	rs28399504	c.1A>G	p.M1V	Rare	None	Ferguson et al. [85]
*5	Exon 9	rs56337013	c.1297G>A	p.R433W	NA	None	Xiao et al. [86] Ibeanu et al. [87]
*6	Exon 3	rs72552267	c.395G>A	p.R132Q	NA	None	Ibeanu et al. [88]
*7	Exon 5	rs72558186	c.819+2T>A	Splicing defect	NA	None	Ibeanu et al. [89]
*8	Exon 3	rs41291556	c.358T>C	p.W120R	Rare	None	Ibeanu et al. [89]
*9	Exon 3	rs17884712	c.431G>A	p.R144H	Rare	Decreased	Blaisdell et al. [90]
*10	Exon 5	rs6413438	c.680C>T	p.P227L	NA	Decreased	Blaisdell et al. [90]
*17		rs12248560	c.-806C>T	Increased expression	Europeans: 0.217 Asians: 0.022 Africans: 0.275	Increased	Sim et al. [91] Rudberg et al. [92] Sibbing et al. [81]

at 4 hours, and poor responder status (Table 25.2) [95]. Similarly, loss-of-function variants in another gene coding a different CYP450 isoenzyme—*CYP2C9*2* (rs1799853) and *CYP2C9*3* (rs1057910)—were significantly associated with lower exposure to the clopidogrel active metabolite, lower inhibition of platelet aggregation at 4 hours, and poor responder status. Concerning prasugrel, the authors did not observe a relationship between the *CYP2C19* or *CYP2C9* polymorphisms and exposure to either the active metabolite of prasugrel or pharmacodynamic response [95]. The association of clopidogrel response variability with the genetic variants of the *CYP2C9* gene was not demonstrated in other studies of healthy subjects or studies of CAD patients on clopidogrel therapy [77,96,97].

In 162 healthy subjects treated with a 300 mg loading dose of clopidogrel, Mega and co-workers reported that

TABLE 25.2 Association of *CYP2C19*2* Loss-of-Function Polymorphisms with Clopidogrel Response Variability in Healthy Subjects

Study	Number of Subjects	Treatment	Functional Evaluation	*CYP2C19*2* Carriers (%)	Association
Hulot et al. [93]	28 Caucasian males	75 mg/day for 7 days	10 μM ADP-induced PA	28.6	Yes
Fontana et al. [94]	94 Caucasians	300 mg + 75 mg/day for 7 days	20 μM ADP-induced PA	27.7	Yes
Brandt et al. [95]	74 Caucasians (57 male/17 female)	300 mg	20 μM ADP-induced PA; active metabolite by LC + MS	25.8	Yes
Mega et al. [96]	162 (130 male/32 female)	300–600 mg + 75 mg/day for 7 days	20 μM ADP-induced PA; active metabolite by LC + MS	34.0	Yes
Kim et al. [98]	24 Koreans	300 mg + 75 mg/day for 7 days	5 μM ADP-induced PA; PCC by LC + MS	66.6 (selected)	Yes
Umemura et al. [99]	47 Japanese	300 mg	20 μM ADP-induced PA; VASPph; active metabolite by LC + MS	61.7	Yes
Shouldiner et al. [97]	429 Amish (214 male/215 female)	300 mg + 75 mg/day for 6 days	20 μM) ADP-induced PA	32.9	Yes

LC + MS—liquid chromatography with tandem mass spectrometry; PA—platelet aggregation; PCC—clopidogrel plasma concentrations; VASPph—vasodilator-stimulated phosphoprotein phosphorylation state of whole blood performed by dual color flow cytometric assay.

carriers of at least one *CYP2C19* loss-of-function allele (30% of the study population) had a relative reduction of 32.4% in plasma exposure to the active metabolite of clopidogrel, as compared with noncarriers (Table 25.2) [96]. Carriers also had an absolute reduction in maximal platelet aggregation in response to clopidogrel that was 9% less than that seen in noncarriers [96].

Kim and co-workers divided 24 Korean subjects into three groups on the basis of their *CYP2C19* genotype: homozygous extensive metabolizers (homoEMs, n = 8), heterozygous EMs (heteroEMs, n = 8), and poor metabolizers (PMs, n = 8) (Table 25.2) [98]. After a single 300 mg loading dose of clopidogrel on day 1, followed by a 75 mg daily maintenance dose on days 2–7, the investigators measured the plasma levels of clopidogrel and assessed the antiplatelet effect. The mean clopidogrel area under the curve (AUC) for PMs was 1.8 and 2.9 times higher than that for heteroEMs and homoEMs, respectively. The mean peak plasma concentration in PMs was 1.8 and 4.7 times higher than that of heteroEMs and homoEMs, respectively. PMs exhibited a significantly lower antiplatelet effect than heteroEMs or homoEMs [98].

In 47 Japanese healthy subjects, Umemura and colleagues determined plasma concentrations of the active metabolite before and at 0.25, 0.5, 1, 2, 4, and 8 hours after treatment with 300 mg clopidogrel (Table 25.2) [99]. Moreover, before and at 1, 4, and 6 hours and after 24 hours, they measured inhibition of platelet aggregation to ADP and VASP, confirming that the *CYP2C19*2* polymorphism is a determinant for active metabolite formation and demonstrating that the pharmacokinetic parameters of the active metabolite (area under the plasma concentration–time curves (AUC); peak concentration (C_{max})) were well correlated with the platelet reactivity index by VASP and inhibition of platelet aggregation by ADP. Their findings suggest that low or no responsiveness to clopidogrel is mainly caused by the lower metabolic formation of the metabolite, but not because of the poor sensitivity of the platelet $P2Y_{12}$ receptor [99].

In an elegant study published in the 2009, known as PAPI (Pharmacogenomics of Antiplatelet Intervention, Shuldiner and colleagues administered clopidogrel—300 mg oral loading dose followed by 75 mg per day for 7 days—to 429 healthy Amish individuals and measured response by *ex vivo* platelet aggregometry (baseline platelet aggregation and 1 hour following the last dose of clopidogrel) (Table 25.2) [97]. They then performed a genome-wide association study (GWAS) followed by the genotyping of the *CYP2C19*2* polymorphism. The presence of relatives in the population allowed the investigators to observe that platelet response to clopidogrel was highly heritable ($h2 = 0.73$; $p < 0.001$). They identified 13 polymorphisms on chromosome 10q24 in the *CYP2C18-CYP2C19-CYP2C9-CYP2C8* cluster associated with diminished clopidogrel response. All of these polymorphisms were in strong linkage disequilibrium with the genetic variant rs12777823, the most significantly associated polymorphism ($p = 1.5 \times 10^{-13}$ at the additive model). rs12777823 was in strong linkage disequilibrium with the *CYP2C19*2* variant, accounting for 12% of the variation in platelet aggregation to ADP ($p = 4.3 \times 10^{-11}$). No other genomic region revealed association signals that met or exceeded genome-wide significance ($p < 1.0 \times 10^{-7}$) [97]. These data suggest that, in this population of healthy subjects, the common loss-of-function *CYP2C19*2* variant may account for most of the association signal detected in the GWAS. Because of the high heritability of platelet response to clopidogrel, other common SNPs not present on the chip, other types of variants (e.g., copy number variants, insertions/deletions), or rare variants with large effect sizes not tagged by the genotyped SNPs may play an important role in determining clopidogrel response variability.

25.3.3 Association of *CYP2C19*2* with High on-Clopidogrel Platelet Reactivity in High-Risk Vascular Patients

In 2007, Giusti and co-workers were the first to study the effect of the *CYP2C19*2* loss-of-function allele on platelet function modulation in ACS patients who had undergone PCI and stent implantation on dual antiplatelet treatment (Table 25.3) [100]. They demonstrated that *CYP2C19*2* is associated with higher platelet aggregability and HPR in 1419 high-risk vascular patients on dual antiplatelet treatment. Carriers of the *CYP2C19*2* allele had significantly higher platelet aggregation values than noncarriers after 2 μM and 10 μM ADP, and after arachidonic acid (AA) stimuli. In particular, homozygous subjects for the *CYP2C19*2* allele had significantly higher platelet aggregation after ADP stimulus (2 μM and 10 μM) than heterozygous carriers. The genotype distribution of *CYP2C19*2* significantly differed between patients with and without HPR evaluated by 10 μM ADP-induced platelet aggregation ≥70%. The prevalence of carriers of the minor allele was higher in patients with HPR by ADP than in patients without HPR by ADP. Similar results were also observed in patients with HPR by AA compared with patients without HPR by AA.

In the multivariate linear regression model—with ADP 2 μM or ADP 10 μM or log(AA) platelet aggregation as the dependent variable and age, gender, hypertension, diabetes mellitus, dyslipidemia, and smoking habit as the independent variables—age, diabetes, and *CYP2C19*2* remained significant and independent risk factors for antiplatelet treatment variability [100]. In that study, Guisti et al. also found no influence on clopidogrel platelet aggregability in ACS patients of previously suspected and studied polymorphisms: the +12G/A IVS10 (rs2242480) polymorphism of the *CYP3A4* gene and the T744C (rs2046934) polymorphism of the *P2Y12* gene [100].

Several other studies confirm that the **2* allele of the *CYP2C19* gene is associated with higher platelet aggregability and HPR in similar and different clinical settings. Table 25.3 lists selected representative studies from the literature on the association of *CYP2C19*2* with on-clopidogrel HPR in high-risk vascular patients.

In 603 patients with non ST-elevation ACS, Frere and colleagues demonstrated that *CYP2C19*2* influences post-treatment platelet reactivity and clopidogrel response assessed by ADP-induced platelet aggregation, VASP phosphorylation index, and ADP-induced P-selectin expression (Table 25.3) [101].

Trenk and colleagues studied and followed 797 patients undergoing elective coronary stent implantation; they showed an association between *CYP2C19*2* and platelet function phenotype of high (>14%) ADP residual platelet aggregation on clopidogrel (Table 25.3) [102].

Geisler and co-workers studied 237 symptomatic CAD patients (both stable angina and ACS). They observed that *CYP2C19*2* carriers showed significantly increased residual platelet aggregation compared with noncarriers (Table 25.3) [103].

In the randomized PRINC (Plavix Response in Coronary Intervention) trial involving 60 patients undergoing elective PCI, Gladding and co-workers measured platelet function using the VerifyNow $P2Y_{12}$ analyzer after a 600 mg or split 1200 mg loading dose and after a 75 or 150 mg daily maintenance dose (Table 25.3) [104]. *CYP2C19 *1*1* homozygotes had greater platelet inhibition 2 hours after a 600 mg dose, compared with platelet inhibition in *CYP2C19*2* or *CYP2C19*4* and *CYP2C19*17* carriers. *CYP2C19*2* or *CYP2C19*4* carriers had greater platelet inhibition with the higher loading dose than with the lower dose at 4 hours, and they responded better to the higher maintenance dose regimen.

The association of the *CYP2C19*2* polymorphism with HPR was observed in Japanese patients by Jinnai and co-workers. They found a higher prevalence of *CYP2C19*2* carriers (56%) compared to patients in Western countries (30%) (Table 25.3) [105].

In 2012, Price and co-workers sought to evaluate the influence of SNPs on the pharmacodynamic effect of high- or standard-dose clopidogrel after PCI. In the GIFT (Genotype Information and Functional Testing) study, the GRAVITAS trial screened patients with platelet function testing (VerifyNow $P2Y_{12}$) after PCI at 12–24 hours, and randomly assigned those with high on-treatment HPR to either a high (150 mg) or a standard (75 mg) dose of clopidogrel [106]. A cohort of patients without HPR was also

TABLE 25.3 Association of *CYP2C19*2* Loss-of-Function Polymorphisms with Clopidogrel Response Variability in High-Risk Vascular Patients

Study	Number of Subjects	Treatment	Functional Evaluation	*CYP2C19*2* Carriers (%)	Association
Giusti et al. [100]	1419 PCI + stent	600 mg LD + 75 mg/day maintaining	10 μM ADP-induced PA	31.4	Yes
Frere et al. [101]	603NSTE	600 mg LD	10 μM ADP-induced PA; VASPph; ADP-induced P-selectin	27.6	Yes
Trenk et al. [102]	797PCI + stent	600 mg LD + 75 mg/day maintaining	5 μM + 20 μM ADP-induced PA; ADP-induced P-selectin; activated GPIIa/IIIb, CD63, CD40L, CD41 (GPIIb)	30.7	Yes
Geisler et al. [103]	237 PCI	600 mg LD + 75 mg/day maintaining	20 μM ADP-induced PA	26.0	Yes
Gladding et al. [104]	60 PCI	600/1200 mg LD + 75/150 mg/day	VerifyNow P2Y$_{12}$		Yes
Jinnai et al. [105]	30 PCI (Japanese)	300 mg LD + 75 mg/day maintaining	5 μM ADP-induced PA	56.0	Yes
Price et al. [106]	1028 PCI + stent	600 mg LD + 75/150 mg/day maintaining	VerifyNow P2Y$_{12}$	36.4	Yes

GPIIa/IIIb—glycoprotein IIa/IIIb (PAC-1); GPIIb—glycoprotein IIb; LD—loading dose; NSTE—non-ST elevation; PA—platelet aggregation; PCI—percutaneous coronary intervention; VASPph—vasodilator-stimulated phosphoprotein phosphorylation state of whole blood performed by dual color flow cytometric assay.

followed. The researchers obtained DNA samples from 1028 patients, who were genotyped for 41 SNPs in 17 genes related to platelet reactivity. They found that *CYP2C19*2* was significantly associated with HPR at 12–24 hours, 30 days, and 6 months after PCI, whereas other candidate SNPs (e.g., *PON1*, *ABCB1*) were not. Carriers of one and two reduced-function *CYP2C19* alleles were significantly more likely to display persistently high HPR at 30 days and 6 months, irrespective of treatment assignment. The portion of the risk of persistently high HPR at 30 days that was attributable to reduced-function *CYP2C19* allele carriage was 5.2% in patients randomly assigned to high-dose clopidogrel. The investigators concluded that CYP2C19, but not PON1 or ABCB1, is a significant determinant of the pharmacodynamic effects of clopidogrel, both early and late after PCI. In patients with high HPR identified by platelet function testing, the *CYP2C19* genotype provides limited incremental information regarding the risk of persistently high reactivity with 150 mg clopidogrel maintenance dosing [106].

25.3.4 Association of *CYP2C19*2* with Major Adverse Cardiovascular Events

In 2009, several groups simultaneously demonstrated, in different clinical settings and at different follow-up times, that the *CYP2C19*2* polymorphism is a significant and independent determinant of MACE and, in particular, of ST (Table 25.4) [96,107–110].

A study conducted in 2008 failed to demonstrate the role of the *CYP2C19*2* loss-of-function polymorphism in the occurrence of MACE or ST. In fact, Trenk and co-workers did not demonstrate a relationship between *CYP2C19*2* polymorphism and increased occurrence of adverse 1-year clinical outcome (death from any cause and MI) of elective PCI with DES or BMS [102]. They showed an association

TABLE 25.4 Association of *CYP2C19*2* Loss-of-Function Polymorphisms with MACE

Study	Number of Subjects	Treatment	Outcome (Follow-up)	CYP2C19 Alleles	Association
Simon et al. [107]	2208 AMI with or without PCI 1535/2208 PCI	300 mg mean LD + 75 mg/day maintaining	Death + nonfatal stroke	Any two **2, *3, *4, *5*	Yes (HR=1.98)
			MI (1 year)		Yes (HR=3.58)
Mega et al. [96]	1477 ACS	300 mg LD + 75 mg/day maintaining	Cardiovascular death + MI + stroke (15 months)	**2*	Yes (HR=3.09)
			ST (15 months)		Yes (HR=1.53)
Collet et al. [108]	259 MI (<45 yrs)	600 mg LD + 75 mg/day maintaining	Death + MI + coronary revascularization (6 months)	**2*	Yes (HR=6.02)
			ST + ischemic events		Yes (HR=3.69)
Giusti et al. [109]	772 PCI + DES	600 mg LD + 75 mg/day maintaining	Cardiovascular death + ST (6 months)	**2*	Yes (OR=2.70)
			ST (6 months)		Yes (OR=3.43)
Sibbing et al. [110]	2485 PCI + stent	600 mg LD + 75 mg/day maintaining	ST (30 days)	**2*	Yes (HR=3.81)

Note: Study results published in 2009. LD—loading dose; PCI—percutaneous coronary intervention; MI—myocardial infarction; AMI—acute myocardial infarction; ACS—acute coronary syndrome; ST—intrastent thrombosis; DES—drug-eluting stent; HR—hazard ratio; OR—odds ratio.

of *CYP2C19*2* with platelet function phenotype, as well as an association of on-clopidogrel HPR with poor clinical outcomes. On the other hand, in their patient population only 280 patients out of 797 had at least one DES. In that group, the percentage of patients with an event was 3.3% (3 out of 90) among carriers of the *CYP2C19*2* allele versus 2.1% (4 out of 190) among wild-type homozygotes [102].

In a study of 2208 patients with acute MI on clopidogrel therapy, Simon and co-workers assessed the association of several allelic variants of genes modulating clopidogrel absorption (*ABCB1*), metabolic activation (*CYP3A5* and *CYP2C19*), and biological activity (*P2Y12* and *ITGB3*) with the risk of death from any cause, nonfatal stroke or MI during 1 year of follow-up [107]. None of the selected polymorphisms in the *CYP3A5*, *P2Y12*, or *ITGB3* genes were associated with the risk of adverse outcomes. Patients with two variant alleles of *ABCB1* (TT genotype) had a higher rate of cardiovascular events at 1 year than those with the *ABCB1* wild-type genotype (CC) (15.5% versus 10.7%; adjusted HR = 1.72, 95% CI 1.20–2.47). Moreover, they demonstrated that patients carrying *CYP2C19* loss-of-function alleles (**2, *3, *4,* or **5*) had a higher rate of subsequent cardiovascular events during 1 year of follow-up (HR 1.98, 95% CI 1.10–3.58) than those who were not. The presence of two variant alleles of *CYP2C19*, but not *ABCB1*, was found to be associated with an increase by a factor of 3.58 (95% CI 1.71–7.51) in the rate of cardiovascular events among the 1535 patients who had undergone PCI during hospitalization, as compared with those who had not (Table 25.4) [107].

Mega and co-workers tested the association of functional genetic variants in CYP genes (*CYP2C19*, *CYP2C9*, *CYP2B6*, *CYP3A5*, *CYP3A4*, and *CYP1A2*), plasma concentrations of active drug metabolites, and platelet inhibition in response to clopidogrel in 162 healthy subjects [96]. Furthermore, they examined the association between these genetic variants and cardiovascular outcomes in a separate cohort of 1477 subjects with ACS who were treated with clopidogrel in the TRITON TIMI 38 trial [96].

In the cohort of 1477 ACS patients who were treated with clopidogrel, Mega and co-workers demonstrated that *CYP2C19*2* carriers had significantly lower levels of the active metabolite of clopidogrel, diminished platelet inhibition, and a higher rate of MACE at 15 months (death from cardiovascular causes, MI, or stroke; HR 1.53, 95% CI1.07–2.19) and an increase (by a factor of 3) in ST risk (HR 3.09, 95% CI 1.19–8.00) [96].

In 259 young patients (<45 years) who had survived a first MI and had been exposed to clopidogrel treatment for at least 1 month, Collet and co-workers assessed whether *CYP2C19*2* affected long-term prognosis in those chronically treated with clopidogrel [108]. They chose as a primary end point a composite of cardiovascular death, nonfatal MI, and urgent revascularization during exposure to clopidogrel. ST was the secondary end point. In this cohort of premature MI patients, the primary end point occurred more frequently in carriers than in noncarriers (15 versus 11 events; HR = 3.69, 95% CI 1.69–8.05, p = 0.0005). Similar results were obtained for the secondary end point; the researchers observed an increase in the risk of ST by a factor of 6 (8 versus 4 events; HR = 6.02, 95% CI 1.81–20.04, p = 0.0009). After multivariable stepwise Cox regression analysis, considering all of the baseline characteristics in the model, if they were significant at p = 0.20 in the univariate model, especially those known to be associated with recurrent acute coronary thrombosis and poor response to clopidogrel (body mass index, smoking status, diabetes mellitus, stent implantation, initial ST-segment elevation MI, and use of PPIs), the *CYP2C19*2* genetic variant was the only independent predictor of cardiovascular events (HR = 4.04, 95% CI 1.81–9.02, p = 0.0006) [108].

Giusti and co-workers studied 772 patients undergoing PCI and stent implantation, who were receiving dual antiplatelet treatment as part of the RECLOSE trial. They evaluated the role of *CYP2C19*2* in drug-eluting ST (primary end point) and the composite feature of cardiac mortality and drug-eluting ST (secondary end point) at 6-month follow-up (Table 25.4) [109]. Patients with ST or the composite of cardiac mortality and ST showed a higher prevalence of *CYP2C19*2* carriers (54.1% versus 31.3%, p = 0.025; 51.7% versus 31.2%; p = 0.020, respectively). In this clinical setting of high-risk vascular patients, the researchers demonstrated that *CYP2C19*2* was an independent risk factor for ST when its effect was evaluated in a model adjusted for on-clopidogrel HPR evaluated by ADP-induced platelet aggregation (ADP-HPR), traditional cardiovascular risk factors, and clinical and procedural risk factors for ST (e.g., total chronic occlusion, multivessel disease, bifurcation lesion, acute MI, previous MI, total stent length, left ventricular ejection fraction) (OR = 3.43, 95% CI 1.01–12.78). Moreover, subjects with contemporaneous *CYP2C19*2* and ADP HPR showed a strong risk of ST or ST and cardiac mortality (OR = 5.79, 95% CI 1.04–39.01 and OR = 11.45, 95% CI 1.84–71.27, respectively) [109].

Sibbing and co-workers investigated the role of the *CYP2C19*2* loss-of-function polymorphism on ST following PCI performed after pretreatment with clopidogrel. The primary end point of the study was the incidence of definite ST within 30 days following PCI. A total of 2485 patients with CAD undergoing PCI were enrolled in this study (Table 25.4) [110].The cumulative 30-day incidence of ST was significantly higher in *CYP2C19*2* allele carriers (*1/*2 heterozygotes or *2/*2 homozygotes) versus *CYP2C19* wild-type homozygotes (*1/*1) (1.5% in *CYP2C19*2* allele carriers versus 0.4% in *CYP2C19* wild-type homozygotes, HR = 3.81, 95% CI 1.45–10.02, p = 0.006, after adjustment for confounding variables). The risk of ST was highest (2.1%) in patients with the *CYP2C19* *2/*2 genotype (p = 0.002) [110].

In contrast to the findings of previous studies, Paré and co-workers showed that, in patients with ACS from the CURE study (formally, Clopidogrel in Unstable Angina to Prevent Recurrent Events) and in those with atrial fibrillation from the ACTIVE A study (formally Fibrillation Clopidogrel Trial with Irbesartan for Prevention of Vascular Events A) *CYP2C19*2* was not associated with an increased risk of MACE [111]. One possible explanation for the discrepancy between these data and previous data can be found in the rate of PCI with stenting. Whereas previous evidence was obtained in cohorts of which the totality, or at least a large proportion, had undergone PCI with or without stenting, only 18% of the patients from CURE had undergone PCI, and only 14.5% had undergone PCI with stent implantation.

Several meta-analyses of clopidogrel pharmacogenetics available in the literature discuss apparently different responses to a similar interrogation [50,112–114]. These apparent discrepancies are used to support doubts about the usefulness of clopidogrel pharmacogenetics. Sofi and co-workers [112] performed a meta-analysis using data available for 7950 patients from six cohort prospective studies who were followed for a time ranging from 6 months to 8.4 years [96,102,107–110]. Their results reinforced the observation that *CYP2C19*2* is associated with increased risk of MACE and, in particular, ST. The summary risk ratios (RR) for included cohort prospective studies showed a significant association between *CYP2C19*2* and increased risk of MACE in the follow-up (RR = 1.80, 95% CI 1.02–3.19, p = 0.04). When studies evaluating ST (n = 4), for a total of 4975 patients were considered, *CYP2C19*2* was associated with an increased risk of ST (RR = 2.82, 95% CI 1.43–5.56, p = 0.0001).

In two larger meta-analyses, one by Hulot and co-workers and the other by Mega and co-workers, it was further demonstrated that the status of carriers of just one loss-of-function *CYP2C19*2* allele is associated with significantly increased risk of MACE, particularly ST [50,113]. With these findings, the *CYP2C19* genetic identifies approximately 30% of the population who may be less likely to gain protection from recurrent ischemic events after PCI, despite treatment with standard doses of clopidogrel.

A further systematic review and meta-analysis of 32 studies by Holmes and colleagues came to different conclusions [114]. By including all studies in which clopidogrel pharmacogenetics were tested relative to cardiovascular events—independently of the fact that they studied high-risk patients undergoing PCI or lower-risk vascular patients—the researchers concluded that CYP2C19*2 carrier status was not clinically relevant for cardiovascular outcomes, with the possible exception of ST [114].

The hazard of drawing conclusions to apply in clinical practice by acquiring data from systematic reviews or meta-analyses—including and considering similar patients with greatly different risks of cardiovascular events and therefore with different benefits gained from clopidogrel treatment—should be underlined [76]. Indeed, Holmes and co-workers paid particular attention to comparator trials of clopidogrel treatment response in which both treatment and comparator group were genotyped (CLARITY-TIMI 28, CURE, ACTIVE-A, CHARISMA) [114]. In these studies, populations consisted of patients who, in large part, had substantial differences in cardiovascular risk in comparison to populations in studies considered in previous meta-analyses: patients receiving fibrinolytic therapy. Of these patients, 50% also had PCI (CLARITY-TIMI 28) with NSTEMI, and 18% had PCI (CURE) with atrial fibrillation (ACTIVE-A). The benefit of clopidogrel in terms of risk reduction for these patients was 11%; for those with documented cardiovascular disease, or risk thereof, clopidogrel showed no benefit over placebo (CHARISMA) [114].

25.3.5 Other Investigated Genetic Variants

Other polymorphisms were studied, but at present clear evidence on their role has not been attainable. In the same issue of the *Lancet*, Mega and co-workers [115] (in a TRITON-TIMI 38 cohort of patients) and Wallentin and co-workers [116] (in a PLATO cohort of patients) evaluated the effect of *CYP2C19*2* and C3435T *ABCB1* (rs1045642) (a synonymous polymorphism in exon 26 substitutes itself for isoleucine at position 1145 in the protein) on antiaggregant treatment efficacy. The two studies confirmed the independent role of *CYP2C19*2* as a determinant of MACE in patients with ACS receiving clopidogrel [115,116], even if a lower impact of the loss-of-function alleles was observed in PLATO compared with TRITON-TIMI 38 or the meta-analyses [50,112,113]. Neither study showed any association between the *CYP2C19* genetic variants and MACE in patients receiving prasugrel or ticagrelor [115,116].

The most interesting novel data from these two studies were those obtained on the C3435T *ABCB1* polymorphism. The Mega and Wallentin articles presented conflicting results here, even though both investigated relatively large populations [115,116]. Mega and co-workers found that patients receiving clopidogrel who had the *ABCB1* 3435 TT genotype were at increased risk of recurrent ischemic events compared with those who had the CT or CC genotypes. The researchers posited that this was most likely because of reduced platelet inhibition in these patients, according to their observations in healthy individuals [115]. In contrast, Wallentin and co-workers observed a higher composite end point of death from cardiovascular-related causes, MI, and stroke in those who had the *ABCB1* 3435 CC genotype (high predicted P-GP expression) and were receiving clopidogrel [116].

When interpreting the results of these two studies, we should consider that several important differences were present, for example, in a clinical setting (only two-thirds of patients were managed invasively in the PLATO cohort), severity of disease, available clinical information, geographic origin, and proportion of platelets naïve to antiplatelet treatment. However, few and contrasting data are available on the potential effect of the *ABCB1* 3435 CT polymorphism on P-GP expression or activity, or on the absorption or metabolism of clopidogrel [27]. In CAD patients who had undergone stent implantation, Bouman and co-workers investigated the clinical relevance of the Q192R (rs662) polymorphism of the gene coding paraoxonase-1 (*PON1*), an enzyme involved in clopidogrel bioactivation [117]. Homozygous *PON1*192 QQ patients showed a considerably higher risk of ST than did homozygous 192 RR patients, reduced PON1 plasma activity, lower plasma concentration of the active metabolite of clopidogrel, and reduced platelet inhibition [117]. Cayla and co-workers performed a sequential analysis of clinical, angiographic, and genetic factors associated with definite early ST [118]. They observed that among the 23 genetic variants investigated in 15 different genes, the significant determinants of early ST were CYP2C19 metabolic status (*CYP2C19*2*, and *CYP2C19*17*) (adjusted OR = 1.99, 95% CI 1.47–2.69), *ABCB1* 3435 TT genotype (adjusted OR = 2.16, 95% CI 1.21–3.88), and *ITGB3 PLA2* carriage (adjusted OR = 0.52, 95% CI 0.28–0.95) [118]. Interestingly, these researchers demonstrated that the combined clinical and genetic model led to a statistically significant increase in the model's discriminatory power compared with the clinical-only model (area under the curve = 0.78 (95% CI 0.73–0.83) versus 0.73 (95% CI 0.67–0.78); $p = .004$) [118].

Recently, Lewis and colleagues measured *ex vivo* platelet aggregation before and after dual antiplatelet therapy in individuals ($n = 565$) from the PAPI study, and conducted a GWAS of drug response [119]. Significant findings were replicated by examining genotype and cardiovascular outcomes in 2 independent aspirin-treated cohorts: 227 PCI patients and 1000 patients in the International Verapamil SR/Trandolapril Study (INVEST) Genetic Substudy (INVEST-GENES). Lewis and colleagues showed that a common polymorphism in the platelet endothelial aggregation receptor-1 (*PEAR1*) gene is a determinant of platelet response and cardiovascular events in patients on aspirin alone or in combination with clopidogrel. In fact, results from the GWAS revealed an association between SNPs on chromosome 1q23 and post-dual antiplatelet therapy platelet aggregation. Further genotyping revealed rs12041331 in *PEAR1* to be most strongly associated with dual antiplatelet therapy response ($p = 7.66 \times 10^{-9}$). Among patients undergoing PCI, A-allele carriers of rs12041331 were more likely to experience a cardiovascular event or death compared with GG homozygotes (HR = 2.62; 95% CI 0.96–7.10, p = 0.059; and HR, 3.97; 95% CI 1.10–14.31, p = 0.035, respectively). In aspirin-treated INVEST-GENES patients, rs12041331 A-allele carriers had significantly increased risk of MI compared with GG homozygotes (OR = 2.03; 95% CI 1.01–4.09, p = 0.048) [119].

25.4 CURRENT STATUS

A large, consistent body of evidence exists on clopidogrel response variability, pharmacogenetics, and alternative therapeutic strategies in high-risk vascular patients. The development of personalized medicine is a concrete opportunity in this clinical setting. The definition of when to perform the genetic test, how many and which genetic variants to include in the diagnostic panel, and what diagnostic tests are available for routine purposes according to the specific context in which the test is carried out are, at present, widely discussed by the scientific community.

25.4.1 Screening, Stratification, and Personalization

In May 2013, the "Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for Cytochrome P450-2C19 (CYP2C19) Genotype and Clopidogrel Therapy: 2013 Update" was published [77]. This update on the clinical use of *CYP2C19* genotype test results for patients requiring antiplatelet therapy includes the appropriate indication for *CYP2C19* genotype-directed antiplatelet therapy, recommendations for specific *CYP2C19* alleles, and additional evidence from an expanded literature review (available at www.pharmgkb.org). In the near future, as a result of the growing ease and availability of genetic testing and other sequencing programs, an increasing number of patients may already know their CYP2C19 genotype status at the time of treatment. Thus, the CPIC guidelines provide guidance on the clinical management of subjects for whom the genotype is available, or for whom the clinician chooses to order a CYP2C19 genotyping test [77].

In July 2013, investigators conducted a prospective randomized controlled trial on antiplatelet therapy tailored on the basis of the CYP2C19 genotype. To the best of our knowledge, clinical outcomes have yet to be reported; therefore, the May 2013 CPIC recommendations are based on the evaluation of evidence available up to that point. Moreover, it is necessary to point out, as do the authors of the guidelines, that therapeutic recommendations are more focused on patients with ACS undergoing PCI. Indeed, the strong evidence for the importance of clopidogrel pharmacogenetics was obtained in clinical settings of high-risk vascular patients. However, this is the clinical setting in which patients draw major benefits—more so than those in other clinical settings—from clopidogrel therapy. Even if, in the clinical setting of ACS patients undergoing PCI and stent implantation, prasugrel and ticagrelor are strongly recommended according to the CPIC update guidelines [6], an optimal individualized antiplatelet treatment should maximize its benefit by reducing the risk of recurrent cardiovascular events while minimizing adverse effects such as bleeding, especially in specific subgroups (e.g., the elderly, women, and those with history of TIA, stroke, or intracranial bleeding). Hence, genotype-directed therapy can identify patients who benefit most from alternative antiplatelet therapy.

In Table 25.5 lists the classifications of subjects' metabolizer status according to the presence of common and rare CYP2C19 polymorphisms.

Concerning treatment with clopidogrel, Table 25.6 summarizes the therapeutic recommendations for antiplatelet therapy based on *CYP2C19* status according to CPIC guidelines: standard dosing of clopidogrel as recommended in the product insert is warranted for patients with ACS and

TABLE 25.5 Metabolizer Status According to CYP2C19 Polymorphisms

Gene	PM	IM	EM	URM
CYP2C19	Subject homozygous/heterozygous compound for **2, *3, *4, *5, *6, *7, *8, *9, *10* loss-of-function alleles	Subject heterozygous for **2, *3, *4, *5, *6, *7, *8, *9, *10* loss-of-function alleles	Subject carrying 2 functional **1* alleles	Subject homozygous/heterozygous for **17* allele

Note: The *CYP2C19*17* allele could balance the effect of the loss-of-function alleles. PM—poor metabolizer; IM—intermediate metabolizer; EM—extensive metabolizer; URM—ultrarapid metabolizer.

TABLE 25.6 **Likely Metabolizer Status (phenotype) Based on CYP2C19 Genotypes and Antiplatelet Therapy Recommendations for ACS/PCI Patients According to CPIC Guidelines**

Phenotype	PM[a]	IM[b]	EM[c]	URM[d]
Genotypes	Subject carrying 2 loss-of-function alleles (**2–*8*) (e.g., **2/*2, *2/*3, *3/*3*)	Subject carrying 1 functional allele **1* and one loss-of-function allele (**2–*8*) or one loss-of-function allele (**2–*8*) and 1 increased-activity allele (**17*) (e.g., **1/*2, *1/*3, *2/*17*)	Subject carrying 2 functional **1* alleles (**1/*1*)	Subject carrying 2 increased-activity alleles (**17/*17*) or 1 functional allele and 1 increased activity allele (**1/*17*)
Implication for clopidogrel	Significantly reduced platelet inhibition; increased RPA; increased risk of MACE	Reduced platelet inhibition; increased RPA; increased risk of MACE	Normal (EM) or increased (URM) platelet inhibition; normal (EM) or decreased (URM) RPA[e]	
Therapeutic recommendations	Alternative antiplatelet therapy (if no contraindication)—e.g., prasugrel, ticagrelor	Alternative antiplatelet therapy (if no contraindication)—e.g., prasugrel, ticagrelor	Clopidogrel label recommended dosage and administration	
Classification of recommendations	Strong	Moderate	Strong	

[a]*PM (poor metabolizer): homozygous variant, mutant, low, or deficient activity (∼2–15% of patients).* [b]*IM (intermediate metabolizer): heterozygote or intermediate activity (∼18–45% of patients).* [c]*EM (extensive metabolizer): homozygous wild-type or normal activity (∼35–50% of patients).* [d]*URM (ultrarapid metabolizer): normal or increased activity (∼5–30% of patients).* [e]*The CYP2C19*17 allele may be associated with increased bleeding risk [81]; RPA—residual platelet aggregation; MACE—major adverse cardiovascular events.*

Source: Scott et al. [77].

PCI who are predicted *CYP2C19* extensive metabolizers (EMs) or ultrarapid metabolizers (UM) (i.e., **1/*1, *1/*17,* and **17/*17*). If genotyping from a CLIA-certified laboratory (i.e., a Clinical Laboratory Improvement Amendments laboratory) identifies a patient as a CYP2C19 PM (i.e., **2/*2*), current literature supports the use of an alternative antiplatelet agent (e.g., prasugrel, ticagrelor) when not clinically contraindicated. The most challenging patients are CYP2C19 intermediate metabolizers (IMs) (e.g., **1/*2,*1/*3, *2/*17*) because they have higher on-treatment HPR on average compared to EMs. ACS and PCI *CYP2C19*2* heterozygotes treated with clopidogrel have increased risks for MACE, including ST. These data support switching to an alternative antiplatelet agent for IMs when not contraindicated. However, given the wide interindividual variability in HPR observed among clopidogrel-treated IMs, clinical judgment should also take into account other factors that may place an IM at increased risk of a cardiovascular (or adverse bleeding) event to most effectively individualize therapy [77].

In addition, although the CPIC guidelines focus on *CYP2C19*2* and *CYP2C19*3*, many clinical genotyping platforms include other variant alleles (**4–*8, *17*) that can alter the interpretation of a patient's predicted metabolizer phenotype [77]. The **4–*8* alleles show strong *in vitro* evidence for complete loss-of-function of the CYP2C19 enzyme. Therefore, when these alleles are identified among patients with ACS and PCI being treated with clopidogrel, they should be considered as influencing both clopidogrel metabolism and clinical outcomes consistent with the **2* loss-of-function allele [77].

Several studies evaluated the effect of increasing clopidogrel loading or maintenance dosing in both healthy subjects and ACS patients. They consistently indicated an improved platelet inhibition among *CYP2C19*2* heterozygotes (IMs) based on *ex vivo* platelet aggregation, but not significant improvement among homozygotes (PMs) [104,106,120–127].

On the other hand, large clinical trials investigating the effect of higher-dose clopidogrel in ACS patients undergoing PCI with on-clopidogrel HPR concluded that adjusting the dosage based on platelet function monitoring alone does not reduce the incidence of death from cardiovascular causes, MI, or ST [18,39,128]. Although we should keep the results of these clinical trials in mind when discussing this issue, we also need to observe that these trials mainly doubled the clopidogrel maintenance dose, which may not be sufficient based on recent studies suggesting that higher clopidogrel doses may be required to achieve adequate platelet inhibition among *CYP2C19*2* heterozygotes [127].

On the basis of this evidence, and because few clinical studies have investigated the eventual improvement of clinical outcomes according to clopidogrel dose adjustment based on CYP2C19 status alone, it is premature to support an increased dosing strategy based on *CYP2C19* genotype at this time. We should wait for future guidelines incorporating results from ongoing clinical trials that are investigating pharmacogenetics or from those that are evaluating the clinical management of antiplatelet therapy for ACS patients undergoing PCI, including platelet function monitoring as well as genetic and nongenetic determinants of clopidogrel response variability.

As can be seen from the available evidence, *CYP2C19* genetic testing has the potential for allowing clinicians to better evaluate and choose the best antiplatelet therapy when considering treatment of patients with ACS undergoing PCI whose genotypes confer a higher risk of MACE. However, at present the lack of randomized clinical trial evidence that *CYP2C19* genotyping improves outcomes must be acknowledged. Furthermore, although *CYP2C19* genotyping is straightforward and reliable when performed in qualified laboratories, as with any laboratory test an additional possible risk to the patient may be caused by an error in genotyping. Because genotypes are lifelong, any such error may have adverse health implications for the patient. On the other hand, commercial tests and tests approved by the CEIVD (European Communities *In Vitro* Diagnostic) or the FDA are available. Some of these can provide genotyping results in a few hours and have point-of-care features that do not require specialized personnel.

25.4.2 Pharmacogenetic Testing

Because *CYP2C19* genetic variants are implicated in inadequate clopidogrel response, the FDA added a box warning on clopidogrel's label about the diminished effectiveness of the drug in patients whose ability to convert it into its active form is impaired because of unfavorable genotypes (Tables 25.5 and 25.6). Besides indicating that some patients may be poor clopidogrel metabolizers because of genetically determined low CYP2C19 activity, the box reports that genetic testing is available to identify patients with the implicated genetic variants and that, in poor metabolizers of the drug, alternative treatment strategies should be considered (see http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm203888). In fact, because it is important to identify high-risk individuals and to improve patient risk stratification, there is growing interest in genetic testing to identify *CYP2C19* genetic variants and other gene polymorphisms related to clopidogrel pharmacogenetics.

The introduction of pharmacogenetic testing in addition to platelet function testing is another tool for clinicians to use in managing differentiated therapeutic strategies and in tailoring antiplatelet therapy to the characteristics of individual patients, thus amplifying the concept of "personalized medicine."

Beside the fact that the clinical benefits of tailored antiaggregant therapies have been demonstrated in prospective clinical trials, the introduction of clopidogrel pharmacogenetic tests into clinical practice may highlight both the need for the rapid availability of results and the need to consider the costs associated with genetic tests [129]. In fact, clopidogrel pharmacogenetic testing results must be available in a very short time before or during the management of ACS patients in the catheterization laboratory in order to improve clinicians' decisions in the early high-risk phase of ACS and its treatment. Moreover, a pharmacogenetic test must be simple to use, fast, and inexpensive. In particular, it must be performed in the catheterization laboratory or at the bedside (point-of-care) to have a real-time effect.

Clopidogrel pharmacogenetic testing is already available in clinical laboratories. However, it can be technically demanding, time-consuming, and expensive, thus limiting its use in the acute clinical care setting. Indeed, in the clinical setting of ACS patients undergoing PCI, it is fundamental that testing be available as early as possible from a nearby specialized clinical laboratory or from a rapid point-of-care platform.

Classical methods of genotyping, such as TaqMan, need experienced staff (molecular biologist/biotechnologist/technician); thus, they are suitable for research purposes but inadequate for providing diagnostic responses in situations of urgency. New, rapid genotyping methods must necessarily be easy to use by unskilled personnel and must provide a rapid turnaround of results at a reasonable cost. Most especially, they must be executable at the patient bedside in order to allow cardiologist to obtain a rapid genetic profile of patients under treatment and to combine this information in an accurate predictive algorithm to achieve truly antiplatelet-personalized care.

Table 25.7 lists CYP2C19 standard and rapid genotyping technologies/assays that are commercially available for clopidogrel pharmacogenetics.

25.4.2.1 Rapid FDA/CEIVD-Approved Clopidogrel Pharmacogenetic Testing

Verigene®

The Verigene system (Nanosphere, Northbrook, IL) is a molecular diagnostics platform for genotyping variants via an automated blood sample-to-results approach. It employs gold nanoparticle technology for genomic detection, making current genomic DNA amplification by polymerase chain reaction (PCR) unnecessary. The specific cartridge for clopidogrel pharmacogenetics testing, the *CYP2C19* Nucleic Acid Test, allows analysis of *CYP2C19*2*, *CYP2C19*3*, and *CYP2C19*17*. As the starting material for the test, fresh

TABLE 25.7 Commercially Available Clopidogrel Pharmacogenetics Genotyping Technologies

Test	Company	FDA/ CEIVD-approved	Time Analysis	Sample Type	CYP2C19 Alleles Investigated	Laboratory Expertise	Test Type
Standard Genotyping							
TaqMan	Life Technologies	Yes	DNA extraction + 1–2 hours	DNA	_*2, *3, *4, *5, *6, *7, *8, *9, *11, *12, *13, *14, *15, *17_	Excellent	RT-PCR TaqMan probes
AmpliChip® CYP450	Roche Molecular Systems	Yes	—	DNA	_*2/*3_	Excellent	Microarray
Rapid Genotyping							
Verigene	Nanosphere	Yes	3 hours	Whole blood	_*2/*3/*17_	Good	Gold nanoparticles
Spartan RX	Spartan Bioscience	Yes	60 minutes	Buccal mucosa cells	_*2/*3_	Good	PCR
i-densy	Arkray	—	90 minutes	Whole blood, oral swabs, purified DNA	_*2/*3_	Very good	QP (quenching Probe)
INFINITI CYP2C19 Assay	Auto-Genomics	Yes	8 hours	Whole blood	_*2/*3/*17_	Very Good	Microarray (ASPE)
eSensor Technology	GenMark Diagnostics	No	DNA extraction + 3–4 hours	DNA	_*2, *3, *4, *5, *6, *7, *8, *9, *10, *13, *17_	Very good	Microarray
Encompass Platform/ Rheonix CARDCon-sumable (IVD)	Rheonix	No	—	Fresh and fixed tissue; whole blood; serum-saliva-urine	_*2/*3_	Very good	—
HyBeacons	Evogen	No	DNA extraction + 35–40 minutes	Saliva; urine, swabs; whole blood	User choice	Excellent	PCR
Tag-It Mutation Detection	Luminex	No	8 hours	DNA	_2, *3, *4, *5, *6, *7, *8_	Excellent	ASPE via bead array
Invader	ThirdWave Technologies	—	3 hours	DNA	User choice	Very good	Enzymatic
ParagonDX/ SmartCycler	Cepheid	Yes	DNA extraction + 20–40 minutes	DNA	User choice	Excellent	RT-PCR

(not frozen) whole venous peripheral blood anticoagulated in ethylenediaminetetraacetic acid (EDTA) is required. The Verigene system comprises two different instruments, the reader (control unit) and the processor (extractor, hybridization, and scanner unit).

The touch-screen interface and barcode scanner in the reader unit allow the user to track samples, monitor test processing, analyze data obtained from the processor unit, and generate reports. The processor performs automated "hands-off" sample preparation, nucleic acid extraction, sonication, and target hybridization on the specific cartridge.

The cartridge (Figure 25.5) is a laboratory consumable holding prepackaged reagents and the microarray slide (a solid substrate). It contains fluidic compartments with reagents ready to use: hybridization, wash solutions, detection probes, gold nanoparticles, and detection reagents (all of which are automatically transferred to the microarray slide while the cartridge is in the processor). The identification of polymorphisms takes place on the cartridge. In order to trace samples, the cartridge is labeled with a barcode identification number. The extraction tray and the tip holder—a single-use plastic holder containing pipette tips and a rubber plug—are placed in the processor, and transfer the sample and reagents during extraction and subsequently the DNA sample to the test cartridge.

Verigene's technology uses gold nanoparticles (Figure 25.5): after DNA extraction and sonication, the genetic sequences of interest are hybridized on the cartridge microarray slide to complementary fixed sequences. Afterwards, the hybridized sequences fixed on the array are incubated, with the gold nanoparticles functionalized with complementary probes to the sequences of interest. Then the unbound nanoparticles are washed and the bound nanoparticles are amplified by a reduction reaction with silver nitrate to produce a visible spot on the array. The reader captures the image of post-silver amplification, and the spot signal intensities are used by the software to genotype samples and to generate the analysis report.

Verigene's workflow is very simple and takes 3 hours. Moreover, the system can be used by operators with modest understanding of molecular biology and genetic diagnostics. Indeed, after the addition of 1 ml of venous peripheral blood anticoagulated in EDTA by the vacutainer tube and loading onto the processor of consumables (cartridge,

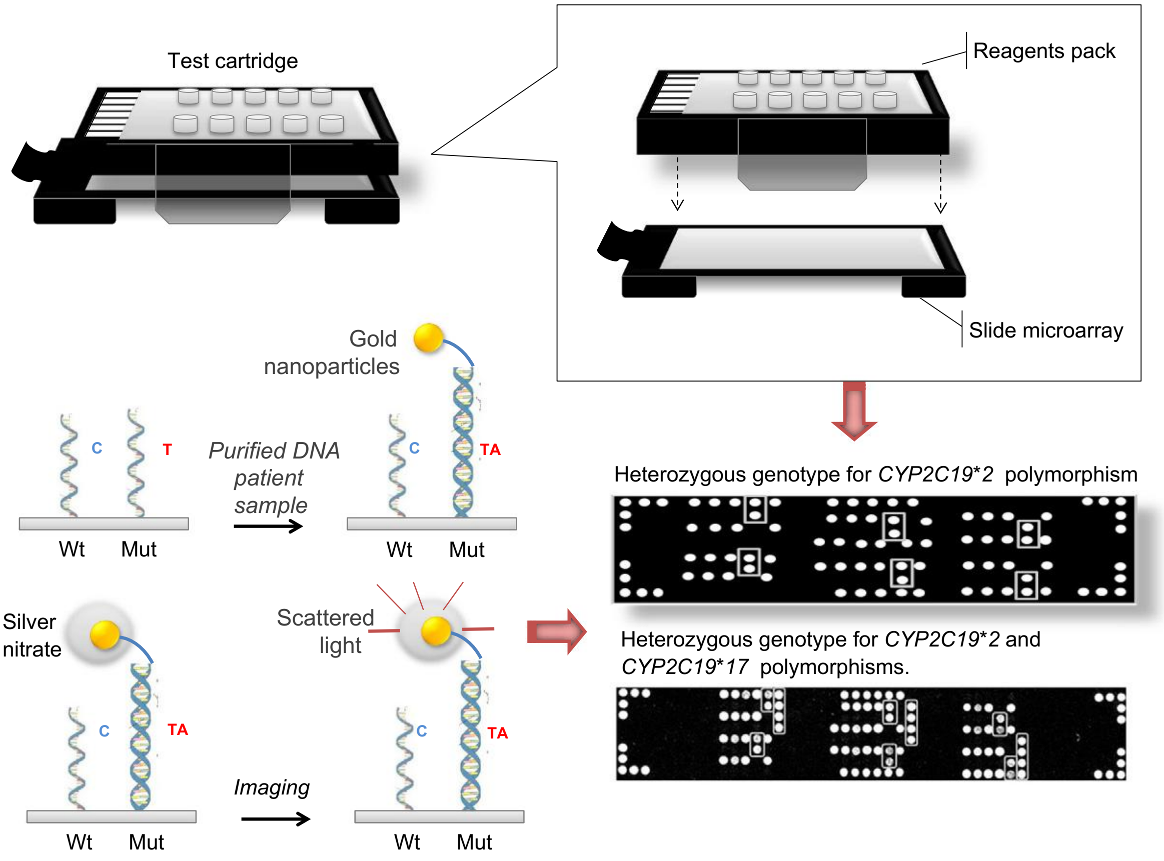

Test cartridge
Reagents pack
Slide microarray
Gold nanoparticles
C
T
Purified DNA patient sample
C
TA
Wt
Mut
Wt
Mut
Heterozygous genotype for *CYP2C19*2* polymorphism
Silver nitrate
Scattered light
C
TA
C
TA
Heterozygous genotype for *CYP2C19*2* and *CYP2C19*17* polymorphisms.
Imaging
Wt
Mut
Wt
Mut

FIGURE 25.5 Microarray and gold nanoparticles used in the Verigene system. The upper drawing represents the structure of the cartridge; the lower part, the technology and the post-silver amplification spots as captured by the analyzer and used by the software for the genotype call. *Source: from Saracini et al.* [129].

extraction tray, and tip holder), the operator can move away and return to move the cartridge in the reader after the 3 hours are up [129].

The accuracy of the Verigene system in identifying *CYP2C19*2* and *CYP2C19*3* was evaluated by Buchan and co-workers [130] using 1286 samples and sequencing validation. The Verigene assay accurately identified homozygous and heterozygous *CYP2C19*2* and *CYP2C19*3* genotypes with a specificity of 100% and a final call rate of 99.7% [130].

Another fundamental feature of rapid genetic testing in clinical practice, besides accuracy and turnaround time, is cost. Saracini and colleagues compared the rapid point-of-care Verigene system with TaqMan, a standard genotyping method [131]. Using both Verigene and TaqMan, they genotyped for *CYP2C19*2*, *CYP2C19*3*, *CYP2C19*4*, *CYP2C19*5*, and *CYP2C19*17* genetic variants in 100 ACS patients, and observed a 100% concordance between the two technologies for all five analyzed polymorphisms. Saracini and colleagues showed that Verigene has a shorter turnaround time than TaqMan; however, they found the cost of its reagents to be higher. Still, staff involvement and relative cost resulted in a higher total cost for TaqMan than for Verigene [131].

Spartan RX

The Spartan RX system (Spartan Biosciences, Ottawa, ON) is a CEIVD-approved portable point-of-care test for fast and accurate analysis of *CYP2C19*2*, *CYP2C19*3*, and *CYP2C19*17*. It is very rapid, with results available in 1 hour, because genomic DNA extraction is not required; instead, DNA is obtained by lysis from cells of the buccal mucosa.

Spartan RX consists of an analyzer, a thermocycler for real-time PCR (with two optical detection channels for analysis of fluorescence emission relative to the specific probes for the two alleles of each polymorphism), and disposable plastic materials and reagents for genotyping (Figure 25.6). The system is connected to a computer with dedicated software for data analysis and management (e.g., sample identification, report preparation).

Spartan RX's consumables are dedicated and are prepackaged in the Sample Collection Kit. Each kit includes two swabs to collect buccal cells and two 0.2 ml tubes filled with all necessary reagents for real-time PCR and allelic discrimination (Figure 25.6). Swab sticks consist of a cap buffer mounted at one end. Cell samples are collected directly on the cap, the swab sticks are then placed in the tube and the tube is inserted in the PCR instrument (Figure 25.6). The Spartan system does not require complex procedures or experienced staff.

Roberts and co-workers performed a prospective randomized, proof-of-concept trial, called RAPID GENE, to assess pharmacogenetic approaches to dual tailored antiplatelet treatment after PCI in ACS patients. Two hundred patients undergoing PCI for ACS or stable angina were *CYP2C19*2*-genotyped by Spartan RX and then randomly assigned according to genotype to a tailored treatment or to a standard treatment (75 mg/day of clopidogrel) [132]. Carriers of the *CYP2C19*2* allele in the tailored branch were treated with 10 mg/day of prasugrel, while noncarriers were treated with 75 mg/day of clopidogrel. Seven days after PCI, the carriers of *CYP2C19*2* treated with prasugrel showed a significantly lower rate of on-treatment HPR compared with *CYP2C19*2* carriers in the standard treatment group. At follow-up, 23 of 91 patients (25%) in the genotype-tailored group and 23 of 96 patients (24%) in the standard treatment group were carriers of at least one *CYP2C19*2* allele. While no *CYP2C19*2* carriers treated with prasugrel showed HPR (defined as $PRU > 234$), 7 of those in the arm treated with 75 mg/day of clopidogrel did show HPR. The study revealed the Spartan RX system's sensitivity to be 100% with a specificity of 99.3% [132].

Saracini and co-workers compared Spartan RX with the classical TaqMan (Life Technologies, Carlsbad, CA) genotyping technology [129]. The aim of their study was to evaluate this commercially available system in terms of accuracy, turnaround time, cost, and management in a diagnostic laboratory. Genotyping was performed for *CYP2C19*2* with the two technologies on 100 ACS patients undergoing PCI. The results obtained by Spartan RX and TaqMan showed a 100% concordance. The study also compared both Verigene and Spartan RX against TaqMan [129]. Spartan RX provided results more quickly (60 minutes); however, its cost was higher and it could genotype only a limited number of polymorphisms contemporaneously. Although the system is potentially upgradable with additional polymorphisms, at present it seems numerically limited.

Verigene is more expensive than TaqMan, but it can be considerably extended (~80 polymorphisms) to a number of additional genetic variants that can be tested at the same time and in the same analysis session. Apparently, TaqMan does not have a relatively longer total turnaround than the other two technologies (292 minutes for Taqman; 71 minutes for Spartan RX; 206 minutes for Verigene). However, compared to Verigene and Spartan RX, Taqman's time to complete depends on how long the operator takes to conduct the analysis. In sum, even if TaqMan's reagent costs are lower compared to the other systems, its manual operator involvement and related costs are higher (Table 25.8) [129].

Extraction and genotyping with real-time PCR and TaqMan can be automated and organized so that pharmacogenetic results are rapidly made available (a few minutes to a few hours) in a molecular biology laboratory equipped for analysis of diagnostic and prognostic genetic factors. However, it is fundamental to have a work platform

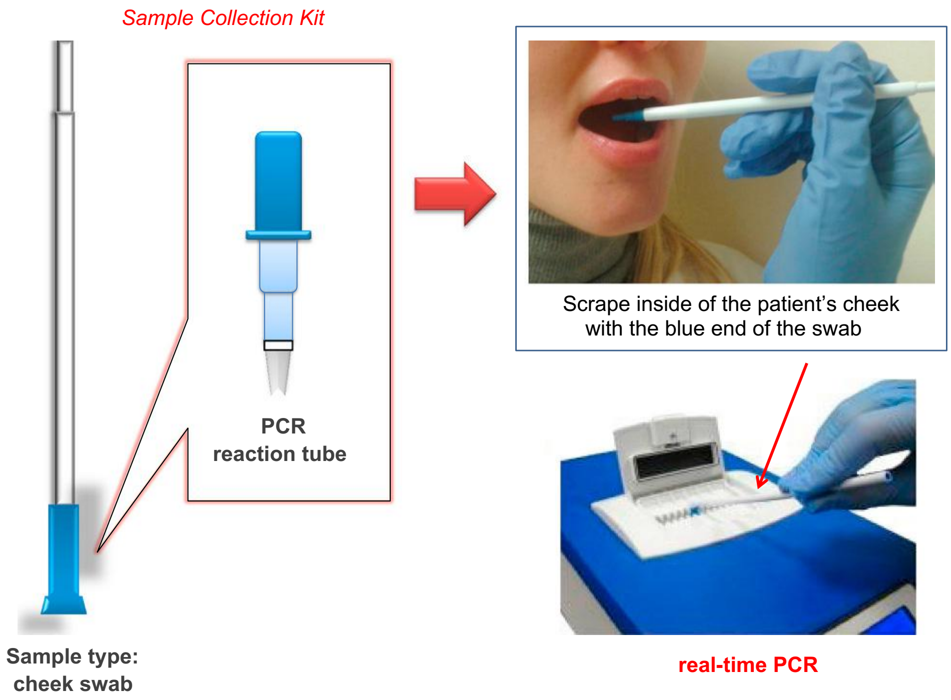


FIGURE 25.6 Real-time PCR-based Spartan RX system.

(for extraction and genotyping) on which the genotyping session can be started at any time of the day without waiting for completion of analyses begun previously. This need for ad hoc work sessions for individual patients makes TaqMan's costs quite high, which is why Taqman is used primarily for research purposes. Verigene and Spartan RX seem to be more compatible with tests carried out by operators unskilled in molecular biology techniques. Moreover, given their characteristics and dimensions, these systems are particularly suitable for genetic testing directly at the patient's bedside—in a catheterization room, for example—for immediate management of therapy for the patient even in situations where a molecular biology/genetics laboratory is not available in the immediate vicinity [129].

TABLE 25.8 Comparative Times and Costs of Genotyping: TaqMan, Verigene, and Spartan RX

	TaqMan	Verigene	Spartan RX
Informatic work plan	3 min	3 min	3 min
Reagents/sample preparation	35 min	20 min	5 min
DNA extraction	200 min	180 min	60 min
DNA quantification	5 min		
Genotyping	10 min + 30 min		
Data analysis	9 min	3 min	3 min
Operator manipulations	Many	Few	Few
Total time per 3 SNPs	62 min + 230 min	6 min + 200 min	6 min + 65 min
Biologist/scientific laboratory technician	€52	€5 biologist/scientific laboratory technician	€5 biologist/scientific laboratory technician
Materials costs	€30	€90	~€100
Total cost	€82	€95	~€105

25.4.2.2 *Other CYP2C19 Genotyping Technologies*

AmpliChip® CYP450

The AmpliChip® CYP450 Test (Roche Molecular Systems, Basel, Switzerland) is the first microarray-based *in vitro* diagnostic test that is FDA approved for clinical use. It allows the detection in the *CYP2D6* and *CYP2C19* genes (Table 25.7) of up to 33 *CYP2D6* alleles and 3 *CYP2C19* alleles (*CYP2C19*1*, *CYP2C19*2*, *CYP2C19*3*).

The AmpliChip test runs on the Affymetrix GeneChip Microarray Instrumentation System (Affymetrix, Santa Clara, CA). Data analysis is performed by the GeneChip Operating Software (GCOS). The main criticisms of this test involve cost (of the order of hundreds of dollars) and time (for DNA extraction and the test itself), and the lack of dedicated high-throughput technology that requires a specialized genetics laboratory and personnel.

Luminex® 100/200™ System and xTAG™

The Luminex 100/200 assay and xTAG P450-2C19 assay (Luminex Molecular Diagnostics, Toronto, ON) are designed to analyze 7 allelic variants: *CYP2C19*2*–*CYP2C19*8* (Table 25.7). They are not currently approved by the FDA. Luminex 100/200 combines multiplex PCR, allele-specific primer extension (ASPE) bead hybridization, and fluorescent detection. It requires manual expertise and is completed in about 8 hours.

INFINITI® CYP2C19

The INFINITI CYP2C19 assay (Autogenomics, Vista, CA) is a CEIVD-approved test for analysis of *CYP2C19*2*, *CYP2C19*3*, and *CYP2C19*17* (Table 25.7). It utilizes the CYP450 2C19+ Intellipac™, the CYP450 2C19+ Amp Mix, and the CYP 450 2C19+ BioFilmChip™ microarray. INFINITI requires the extraction of genomic DNA from whole blood. It is based on allele ASPE technology and quantification of generated fluorescent signals. After DNA extraction, all of INFINITIs other steps can be automated by the INFINITI Analyzer (Autogenomics, Vista, CA). Gladding and co-workers used INFINITI in their dose-escalation study to suggest that increasing the dose of clopidogrel in patients with nonresponder polymorphisms increases antiplatelet response [121].

I-Densy™ IS-5320

A novel rapid platform for polymorphism analysis is i-densy IS-5320 (Arkray, Kyoto, Japan), which in a single system combines sample pretreatment, amplification, and signal detection (Table 25.7), allowing results to be obtained in about 90 minutes. i-densy does not require DNA extraction and is suitable for genotyping from whole blood, purified nucleic acids, and oral swabs. The system and its reagents (Compact Reagent Packs) are commercially available for analysis and allow genotyping of *CYP2C19*2* and *CYP2C19*3*. The technique is based on a quenching-probe (QP) system as the principle tool of SNP detection.

eSensor®

The eSensor 2C19 Genotyping Test (Genmark, Carlsbad, CA) is a DNA microarray system based on electrochemical and hybridization detection of genotypes, designed for multiplex testing of the *CYP2C19*2*–**10*, *CYP2C19*13*, and *CYP2C19*17* polymorphisms (Table 25.7). It is commercially available but has not been cleared or approved for marketing by the FDA. eSensor's results are available in ≈3–4 hours after DNA extraction, and the assay requires a very small amount of DNA concentration (0.05 ng/μl). DNA extraction and PCR are necessary prior to genotyping; the assay can process the sample in 3 hours, requiring 40 minutes of hands-on time.

Encompass MDx™ and Rheonix CARD®

The Encompass MDx™ platform (Rheonix®, New York, NY) is an open system for research only and has not been approved by the FDA for IVD use (Table 25.7). It is fully automated and must be used with Rheonix CARD® consumables. Rheonix Plavix PGx CARD was designed to detect *CYP2C19*2* and *CYP2C19*3*, starting from different sample types such as fresh tissue, formalin-fixed tissue, whole blood, serum, saliva, and swabs. DNA is subjected to multiplex PCR; amplicons are then denatured and delivered to the primer extension reactor, and annealing occurs between the denatured amplicon strands and the immobilized "primers." Solid-phase primer extension is initiated, resulting in incorporation of biotinylated dUTP in the extended "primer" strand on the surface. Primer-extended products are detected via incubation with streptavidin-conjugated HRP and substrate.

25.4.2.3 *Other Genotyping Technologies*

Other available genotyping technologies include Invader® Assay (Third Wave Technologies, Madison, WI), HyBeacons® Assay (Evogen, Kansas City, MO), and ParagonDX/SmartCycler (Chepeid, Sunnyvale, CA) (Table 25.7). Invader is a proprietary system for genetic analysis that is similar to PCR. HyBeacons High Definition PCR (HD PCR) utilizes fluorophore-labeled oligonucleotides to detect and identify specific nucleotide sequences post-amplification. Hybeacon's probes can be used in combination with direct analysis of biological samples of different types and thermal cyclers, such as LightCycler (Roche Molecular Systems, Basel, Switzerland), to perform genotyping of SNPs for many samples in 35–40 minutes. Smart Cycler performs sample analysis in 20–40 minutes after DNA extraction. It is based on real-time PCR amplification of target genomic DNA and simultaneous detection of SNPs.

25.5 CONCLUSIONS AND FUTURE DIRECTIONS

Even though guidelines for clopidogrel pharmacogenetics, as well as on-clopidogrel platelet function evaluation, have been the subject of much research [20,77], still unresolved is the personalization of antiplatelet therapeutic strategies aimed at reducing ischemic events without increasing bleeding risk.

We believe that a clear definition of clopidogrel pharmacogenetics' usefulness (e.g., regarding clinical setting(s); polymorphisms to be included in the genetic profile; length of testing; use of genetic profiling alone or with algorithms for other nongenetic determinants) is particularly important in the current clinical context, where alternative drugs with different characteristics and costs are available. Indeed, a generic version of clopidogrel is now available; the patent for prasugrel expired in 2012; and the patent for ticagrelor will expire in 2018. The opportunity to choose the most efficient and economic pharmacologic strategy is crucial for all patients, of course, but it is particularly important in healthcare systems in which expenditures are totally borne by the patient [76]. Indeed, prasugrel and ticagrelor have been shown to be more potent platelet inhibitors and to have a favorable benefit-to-risk ratio compared with clopidogrel in high-risk patients, but this comes at the expense of shifting the balance toward an increased bleeding risk especially in some patient categories (the elderly, women, and those with history of TIA, stroke, or intracranial bleeding) [6,29,31,76]. Moreover, other antiaggregant drugs such as cilostazol and cangrelor are available and might be considered in the management of high-risk vascular patients.

A large body of evidence highlights an urgent need for prospective studies to clarify whether personalized therapeutic strategies will improve outcomes in high-risk patients with vascular disease. It may not be the best strategy to choose the "superior" drug solely on the basis of the results of large clinical trials in which patients with different risk profiles are considered equal. The challenge for the present and for the near future is to choose not the lesser evil, but the greater good. Overcoming this challenge will be possible only by identifying the antiplatelet therapeutic approach that warrants the higher benefit/risk ratio taking into consideration individual patient characteristics [27].

Prospective studies that evaluate antiplatelet treatments tailored to individual patient characteristics—genetic profile, residual platelet reactivity, drug–drug interactions, and traditional and procedural risk factors—are urgently needed. The results of these studies may allow development of algorithms that take into account and weigh genetic or acquired risk factors for cardiovascular complications, and that help clinicians to better identify therapeutic strategies that will provide the greatest benefit for each patient.

On the other hand, evidence from the literature underlines the need for continuing investigation of as yet unidentified genetic and nongenetic factors that contribute to clopidogrel variability, on-treatment HPR, and ultimately poor outcomes [133,134]. Indeed, Marcucci and co-workers showed that on-clopidogrel HPR by ADP is an independent determinant of MACE in good clopidogrel metabolizers (i.e., noncarriers of *CYP2C19*2*). This suggests that both phenotype and genotype are clinically relevant in evaluation of clopidogrel's antiplatelet effect and in the prognostic stratification of ACS patients. It also suggests that additional factors may be implicated [133]. Recently, Frelinger and co-workers found that clopidogrel pharmacokinetics and pharmacodynamics vary widely in healthy subjects, despite rigorous exclusion or control of known disease, polymorphisms (*CYP2C19*, *CYP3A5*, *ABCB1*, *PON1*), noncompliance, co-medications, diet, smoking, alcohol, demographics, and pretreatment platelet hyper-reactivity. As yet unidentified factors also contribute to on-treatment HPR [134].

In general, the pharmacogenetics of clopidogrel in the treatment of cardiovascular disease and that of other drugs in the treatment of noncardiovascular disease represent an important model for identifying strategies and solutions that will improve healthcare. Consider the following:

- Information is available and consistent regarding the impact of pharmacogenetics, particularly on the inadequate efficacy of clopidogrel.
- Available information is consistent regarding *CYP2C19*2*'s association with poor outcomes in the clinical setting of ACS patients undergoing PCI and stent implantation.
- Preliminary data on the possible usefulness of pharmacogenetics testing for clopidogrel are available.

Such convincing evidence has induced the FDA to place a box on clopidogrel's label advising clinicians of possible inadequate responsiveness in genetically determined patients.

Clopidogrel pharmacogenetics represents a sturdy test bench for identifing critical issues that must be addressed, by both the scientific community (clinicians, molecular biologists, bioinformaticians) and industry. Indeed, there are major concerns with clopidogrel pharmacogenetics in the routine management of ACS patients. These involve economic issues and availability of genetic results. To respond to the need for accurate, rapid point-of-care systems, several FDA- or CEIVD-approved solutions have been developed and others are in preliminary development. Even if the costs of commercially available systems are not comparable with those of genotyping systems for research purposes only, their wider diffusion may consistently reduce such costs in the near future. Several rapid systems for genetic testing are now commercially available, guaranteeing results in 1–3 hours. These are point-of-care tests that can be carried out by unskilled personnel and that cost about €100 ($138) (which will be amortized by the reduction in social and economic costs of managing clinical complications in these high-risk patients). In fact, in appropriate clinical settings a reduction in major adverse outcomes may be achieved by a tailored pharmacological strategy in a single patient, and this may result in reduced costs of managing clinical complications. If so, the cost of the pharmacogenetic test is absolutely justified. Moreover, several commercially available systems are truly point-of-care, meaning that pharmacogenetic testing can be performed at the bedside. This too may result in improved management of patients in any phase of the disease. The committed involvement of the scientific community and industry in solving the critical issue of clopidogrel pharmacogenetics represents a virtuous process that has furnished the technology and expertise needed to address future pharmacogenetic challenges [129].

Another, not insignificant, challenge to be overcome is clinicians' lack of confidence in genetic test results, not only with regard to clopidogrel pharmacogenetics but in regard to the future of pharmacogenetics and stratified medicine generally.

Scientific knowledge deriving from genetics, pharmacogenetics, transcriptomics, metabolomics, microbiome, proteomics, and imaging is rapidly increasing. The major challenges now and in the future are the translation and integration of this mass of information in the canon of knowledge of disease and patient management. Three major issues to be addressed are the development of adequate bioinformatics approaches, the continuous multidisciplinary training of all actors in the process (biologists, medical specialists, medical geneticist, bioinformatics experts, etc.), and the development of diagnosis and treatment paths guided by the principles of stratified medicine.

ABBREVIATIONS

AA	arachidonic acid
ABCB1	ATP-binding cassette, subfamily B (MDR/TAP), member 1
ACS	acute coronary syndromes
ADAMTS13	a disintegrin and metalloproteinase with a thrombospondin type 1 motif, member 13
ADP	adenosine diphosphate
ASPE	Allele Specific Primer Extension
BMS	bare metal stent
CAD	coronary artery disease
CEIVD	European Communities *In Vitro* Diagnostic
COX	cyclooxygenase
CYP450	Cytochrome P450
CYP1A2	cytochrome P450, family 1, subfamily A, polypeptide 2
CYP2B6	cytochrome P450, family 2, subfamily B, polypeptide 6 .
CYP2C19	CYP2C19 cytochrome P450, family 2, subfamily C, polypeptide 19
CYP2C9	CYP2C9 cytochrome P450, family 2, subfamily C, polypeptide 9
CYP3A4	CYP3A4 cytochrome P450, family 3, subfamily A, polypeptide 4
DES	drug eluting stent
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EM	extensive metabolizer
eNOS	endothelial nitric oxide synthase
FDA	Food and Drug Administration
GPIa	glycoprotein Ia
GPIb/V/IX	glycoprotein Ib/V/IX complex
GP IIb/IIIa	glycoprotein IIb/IIIa
GPVI	glycoprotein VI
HPR	high platelet reactivity
IM	intermediate metabolizer
MACE	major adverse cardiovascular events
MI	myocardial infarction
NSTEMI	Non ST elevation myocardial infarction
$P2Y_{12}$	Purinergic receptor P2Y, G-protein coupled, 12
PAD	peripheral arterial disease
PCI	percutaneous coronary intervention
PCR	Polymerase chain reaction
PM	poor metabolizer
PPIs	Proton pump inhibitors
PRP LTA	light transmission aggregometry on platelet rich plasma
SELP	P-selectin
SNP	single nucleotide polymorphism
STEMI=	ST elevation myocardial infarction
TIA	transient ischemic attack
TxA2	thromboxane A2
UA	unstable angina
ULvWF	ultra large von Willebrand factor
UM or URM	ultrarapid metabolizer
VASP	platelet vasodilator-stimulated phosphoprotein
vWF	von Willebrand factor

REFERENCES

[1] Davì G, Patrono C. Platelet activation and atherothrombosis. N Engl J Med 2007;357(24):2482–94.

[2] Eikelboom JW, Hirsh J, Spencer FA, Baglin TP, Weitz JI. Antiplatelet drugs: antithrombotic therapy and prevention of thrombosis, 9th ed: American college of chest physicians evidence-based clinical practice guidelines. Chest 2012;141(2 Suppl.):e89S–e119S.

[3] Kaplan ZS, Jackson SP. The role of platelets in atherothrombosis. Hematology Am Soc Hematol Educ Program 2011;2011:51–61.

[4] Fintel DJ. Oral antiplatelet therapy for atherothrombotic disease: overview of current and emerging treatment options. Vasc Health Risk Manag 2012;8:77–89.

[5] Picker SM. Platelet function in ischemic heart disease. J Cardiovasc Pharmacol 2013;61(2):166–74.

[6] Marcucci R, Cenci C, Cioni G, Lombardi A, Giusti B, Gensini GF. Antiplatelets in acute coronary syndrome: personal perspectives. Expert Rev Cardiovasc Ther 2012;10(12):1487–96.

[7] Vandvik PO, Lincoff AM, Gore JM, Gutterman DD, Sonnenberg FA, Alonso-Coello P, et al. American College of Chest Physicians. Primary and secondary prevention of cardiovascular disease: antithrombotic therapy and prevention of thrombosis, 9th ed: American college of chest physicians evidence-based clinical practice guidelines. Chest 2012;141(2 Suppl.):e637S–68S.

[8] Anderson JL, Adams CD, Antman EM, Bridges CR, Califf RM, Casey Jr DE, et al. 2012 ACCF/AHA focused update incorporated into the ACCF/AHA 2007 guidelines for the management of patients with unstable angina/non-ST-elevation myocardial infarction: a report of the American college of cardiology foundation/ American heart association task force on practice guidelines. J Am Coll Cardiol 2013;61(23):e179–347.

[9] Giusti B, Gori AM, Marcucci R, Saracini C, Vestrini A, Abbate R. Determinants to optimize response to clopidogrel in acute coronary syndrome. Pharmgenomics Pers Med 2010;3:33–50.

[10] Matetzky S, Shenkman B, Guetta V, Shechter M, Beinart R, Goldenberg I, et al. Clopidogrel resistance is associated with increased risk of recurrent atherothrombotic events in patients with acutemyocardial infarction. Circulation 2004;109(25):3171–5.

[11] Buonamici P, Marcucci R, Migliorini A, Gensini GF, Santini A, Paniccia R, et al. Impact of platelet reactivity after clopidogrel administration on drug-eluting stent thrombosis. J Am Coll Cardiol 2007;49(24):2312–7.

[12] Gori AM, Marcucci R, Migliorini A, Valenti R, Moschi G, Paniccia R, et al. Incidence and clinical impact of dual nonresponsiveness to aspirin and clopidogrel in patients with drug-eluting stents. J Am Coll Cardiol 2008;52(9):734–9.

[13] Price MJ, Endemann S, Gollapudi RR, Valencia R, Stinis CT, Levisay JP, et al. Prognostic significance of post-clopidogrel platelet reactivity assessed by a point-of-care assay onthrombotic events after drug-eluting stent implantation. Eur Heart J 2008;29(8):992–1000.

[14] Marcucci R, Gori AM, Paniccia R, Giusti B, Valente S, Giglioli C, et al. Cardiovascular death and nonfatal myocardial infarction in acute coronary syndrome patients receiving coronary stenting are predicted by residual platelet reactivity to ADP detected by a point-of-careassay: a 12-month follow-up. Circulation 2009;119(2):237–42.

[15] Sibbing D, Byrne RA, Bernlochner I, Kastrati A. High platelet reactivity and clinical outcome – fact and fiction. Thromb Haemost 2011;106(2):191–202.

[16] Sofi F, Marcucci R, Gori AM, Giusti B, Abbate R, Gensini GF. Clopidogrel non-responsiveness and risk of cardiovascular morbidity. An updated meta-analysis. Thromb Haemost 2010;103(4):841–8.

[17] Aradi D, Komócsi A, Vorobcsuk A, Rideg O, Tokés-Füzesi M, Magyarlaki T, et al. Prognostic significance of high on-clopidogrel platelet reactivity after percutaneous coronary intervention: systematic review and meta-analysis. Am Heart J 2010;160(3):543–51.

[18] Parodi G, Marcucci R, Valenti R, Gori AM, Migliorini A, Giusti B, et al. High residual platelet reactivity after clopidogrel loading and long-term cardiovascular events among patients with acute coronary syndromes undergoing PCI. JAMA 2011;306(11):1215–23.

[19] Aradi D, Komócsi A, Vorobcsuk A, Serebruany VL. Impact of clopidogrel and potent P2Y 12-inhibitors on mortality and stroke in patients with acute coronary syndrome or undergoing percutaneous coronary intervention: a systematic review and meta-analysis. Thromb Haemost 2013;109(1):93–101.

[20] Bonello L, Tantry US, Marcucci R, Blindt R, Angiolillo DJ, Becker R, et al. Working Group on High On-Treatment Platelet Reactivity. Consensus and future directions on the definition of high on-treatment platelet reactivity to adenosine diphosphate. J Am Coll Cardiol 2010;56(12):919–33.

[21] Pereillo JM, Maftouh M, Andrieu A, Uzabiaga MF, Fedeli O, Savi P, et al. Structure and stereochemistry of the active metabolite of clopidogrel. Drug Metab Dispos 2002;30(11):1288–95.

[22] Savi P, Zachayus JL, Delesque-Touchard N, Labouret C, Hervé C, Uzabiaga MF, et al. The active metabolite of Clopidogrel disrupts P2Y12 receptor oligomers and partitions them out of lipid rafts. Proc Natl Acad Sci USA 2006;103(29):11069–74.

[23] Giusti B, Gori AM, Marcucci R, Abbate R. Relation of CYP2C19 loss-of-function polymorphism to the occurrence of stent thrombosis. Expert Opin Drug Metab Toxicol 2010;6(4):393–407.

[24] Kazui M, Nishiya Y, Ishizuka T, Hagihara K, Farid NA, Okazaki O, et al. Identification of the human cytochrome P450 enzymes involved in the two oxidative steps in the bioactivation of clopidogrel to its pharmacologically active metabolite. Drug Metab Dispos 2010;38(1):92–9.

[25] Ding Z, Kim S, Dorsam RT, Jin J, Kunapuli SP. Inactivation of the human P2Y12 receptor by thiol reagents requires interaction with both extracellular cysteine residues, Cys17 and Cys270. Blood 2003;101(10):3908–14.

[26] Farid NA, Smith RL, Gillespie TA, Rash TJ, Blair PE, Kurihara A, et al. The disposition of prasugrel, a novel thienopyridine, in humans. Drug Metab Dispos 2007;35(7):1096–104.

[27] Giusti B, Abbate R. Response to antiplatelet treatment: from genes to outcome. Lancet 2010;376(9749):1278–81.

[28] Gurbel PA, Tantry US. Prasugrel, a third generation thienopyridine and potent platelet inhibitor. Curr Opin Investig Drugs 2008;9(3):324–36.

[29] Wiviott SD, Braunwald E, McCabe CH, Montalescot G, Ruzyllo W, Gottlieb S, et al. TRITON-TIMI 38 Investigators. Prasugrel versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 2007;357(20):2001–15.

[30] VAN Giezen JJ, Nilsson L, Berntsson P, Wissing BM, Giordanetto F, Tomlinson W, et al. Ticagrelor binds to human P2Y(12) independently from ADP but antagonizes ADP-induced receptor signaling and platelet aggregation. J Thromb Haemost 2009;7(9):1556–65.

[31] Wallentin L, Becker RC, Budaj A, Cannon CP, Emanuelsson H, Held C, et al. PLATO Investigators. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 2009;361(11):1045–57.

[32] Gurbel PA, Tantry US. Clopidogrel response variability and the advent of personalised antiplatelet therapy. A bench to bedside journey. Thromb Haemost 2011;106(2):265–71.

[33] Paniccia R, Antonucci E, Gori AM, Marcucci R, Poli S, Romano E, et al. Comparison of different methods to evaluate the effect of aspirin on platelet function in high-risk patients with ischemic heart disease receiving dual antiplatelet treatment. Am J Clin Pathol 2007;128(1):143–9.

[34] Paniccia R, Antonucci E, Maggini N, Miranda M, Romano E, Gori AM, et al. Light transmittance aggregometry induced by different concentrations of adenosine diphosphate to monitor clopidogrel therapy: a methodological study. Ther Drug Monit 2011;33(1):94–8.

[35] Paniccia R, Antonucci E, Gori AM, Marcucci R, Giglioli C, Antoniucci D, et al. Different methodologies for evaluating the effect of clopidogrel on platelet function in high-risk coronary artery disease patients. J Thromb Haemost 2007;5(9):1839–47.

[36] Paniccia R, Antonucci E, Maggini N, Miranda M, Gori AM, Marcucci R, et al. Comparison of methods for monitoring residual platelet reactivity after clopidogrel by point-of-care tests on whole blood in high-risk patients. Thromb Haemost 2010;104(2):287–92.

[37] Breet NJ, van Werkum JW, Bouman HJ, Kelder JC, Ruven HJ, Bal ET, et al. Comparison of platelet function tests in predicting clinical outcome in patients undergoing coronary stent implantation. JAMA 2010;303(8):754–62.

[38] Patti G, Nusca A, Mangiacapra F, Gatto L, D'Ambrosio A, Di Sciascio G. Point-of-care measurement of clopidogrel responsiveness

predicts clinical outcome in patients undergoing percutaneous coronary intervention results of the ARMYDA-PRO (Antiplatelet therapy for Reduction of MYocardial Damage during Angioplasty-Platelet Reactivity Predicts Outcome) study. J Am Coll Cardiol 2008;52(14):1128–33.

[39] Price MJ, Berger PB, Teirstein PS, Tanguay JF, Angiolillo DJ, Spriggs D, et al. GRAVITAS Investigators. Standard- vs high-dose clopidogrel based on platelet function testing after percutaneous coronary intervention: the GRAVITAS randomized trial. JAMA 2011;305(11):1097–105.

[40] Kuchulakanti PK, Chu WW, Torguson R, Ohlmann P, Rha SW, Clavijo LC, et al. Correlates and long-term outcomes of angiographically proven stent thrombosis with sirolimus- and paclitaxel-eluting stents. Circulation 2006;113(8):1108–13.

[41] Hamann GF, Weimar C, Glahn J, Busse O, Diener HC German Stroke Data Bank. Adherence to secondary stroke prevention strategies–results from the German Stroke Data Bank. Cerebrovasc Dis 2003;15(4):282–8.

[42] Serebruany V, Cherala G, Williams C, Surigin S, Booze C, Kuliczkowski W, et al. Association of platelet responsiveness with clopidogrel metabolism: role of compliance in the assessment of “resistance”. Am Heart J 2009;158(6):925–32.

[43] Airoldi F, Colombo A, Morici N, Latib A, Cosgrave J, Buellesfeld L, et al. Incidence and predictors of drug-eluting stent thrombosis during and after discontinuation of thienopyridine treatment. Circulation 2007;116(7):745–54.

[44] Kimura T, Morimoto T, Nakagawa Y, Tamura T, Kadota K, Yasumoto H, et al. Cypher Registry Investigators. Antiplatelet therapy and stent thrombosis after sirolimus-eluting stent implantation. Circulation 2009;119(7):987–95.

[45] Ferreira-González I, Marsal JR, Ribera A, Permanyer-Miralda G, García-Del Blanco B, Martí G, et al. Double antiplatelet therapy after drug-eluting stent implantation: risk associated with discontinuation within the first year. J Am Coll Cardiol 2012;60(15):1333–9.

[46] Park Y, Jeong YH, Tantry US, Ahn JH, Kwon TJ, Park JR, et al. Accelerated platelet inhibition by switching from atorvastatin to a non-CYP3A4-metabolized statin in patients with high platelet reactivity (ACCEL-STATIN) study. Eur Heart J 2012;33(17):2151–62.

[47] Leoncini M, Toso A, Maioli M, Angiolillo DJ, Giusti B, Marcucci R, et al. Pharmacodynamic effects of adjunctive high dose atorvastatin on double dose clopidogrel in patients with high on-treatment platelet reactivity depending on diabetes mellitus status. J Thromb Thrombolysis 2013 [Epub ahead of print].

[48] Leoncini M, Toso A, Maioli M, Angiolillo DJ, Giusti B, Marcucci R, et al. High-dose atorvastatin on the pharmacodynamic effects of double-dose clopidogrel in patients undergoing percutaneous coronary interventions: the ACHIDO (Atorvastatin and Clopidogrel HIgh DOse in stable patients with residual high platelet activity) study. JACC Cardiovasc Interv 2013;6(2):169–79.

[49] Siller-Matula JM, Jilma B, Schrör K, Christ G, Huber K. Effect of proton pump inhibitors on clinical outcome in patients treated with clopidogrel: a systematic review and meta-analysis. J Thromb Haemost 2010;8(12):2624–41.

[50] Hulot JS, Collet JP, Silvain J, Pena A, Bellemain-Appaix A, Barthélémy O, et al. Cardiovascular risk in clopidogrel-treated patients according to cytochrome P450 2C19*2 loss-of-function allele or proton pump inhibitor coadministration: a systematic meta-analysis. J Am Coll Cardiol 2010;56(2):134–43.

[51] Huang B, Huang Y, Li Y, Yao H, Jing X, Huang H, et al. Adverse cardiovascular effects of concomitant use of proton pump inhibitors and clopidogrel in patients with coronary artery disease: a systematic review and meta-analysis. Arch Med Res 2012;43(3):212–24.

[52] Gilard M, Arnaud B, Cornily JC, Le Gal G, Lacut K, Le Calvez G, et al. Influence of omeprazole on the antiplatelet action of clopidogrel associated with aspirin: the randomized, double-blind OCLA (Omeprazole CLopidogrel Aspirin) study. J Am Coll Cardiol 2008;51(3):256–60.

[53] Cuisset T, Frere C, Quilici J, Poyet R, Gaborit B, Bali L, et al. Comparison of omeprazole and pantoprazole influence on a high 150-mg clopidogrel maintenance dose the PACA (Proton Pump Inhibitors and Clopidogrel Association) prospective randomized study. J Am Coll Cardiol 2009;54(13):1149–53.

[54] Juurlink DN, Gomes T, Ko DT, Szmitko PE, Austin PC, Tu JV, et al. A population-based study of the drug interaction between proton pump inhibitors and clopidogrel. CMAJ 2009;180(7):713–8.

[55] Ho PM, Maddox TM, Wang L, Fihn SD, Jesse RL, Peterson ED, et al. Risk of adverse outcomes associated with concomitant use of clopidogrel and proton pump inhibitors following acute coronary syndrome. JAMA 2009;301(9):937–44.

[56] O’Donoghue ML, Braunwald E, Antman EM, Murphy SA, Bates ER, Rozenman Y, et al. Pharmacodynamic effect and clinical efficacy of clopidogrel and prasugrel with or without a proton-pump inhibitor: an analysis of two randomised trials. Lancet 2009;374(9694):989–97.

[57] Parri MS, Gianetti J, Dushpanova A, Della Pina F, Saracini C, Marcucci R, et al. Pantoprazole significantly interferes with antiplatelet effect of clopidogrel: results of a pilot randomized trial. Int J Cardiol 2013;167(5):2177–81.

[58] Goodman SG, Clare R, Pieper KS, Nicolau JC, Storey RF, Cantor WJ, et al. Platelet Inhibition and Patient Outcomes Trial Investigators. Association of proton pump inhibitor use on cardiovascular outcomes with clopidogrel and ticagrelor: insights from the platelet inhibition and patient outcomes trial. Circulation 2012;125(8):978–86.

[59] Gremmel T, Steiner S, Seidinger D, Koppensteiner R, Panzer S, Kopp CW. Adenosine diphosphate-inducible platelet reactivity shows a pronounced age dependency in the initial phase of antiplatelet therapy with clopidogrel. J Thromb Haemost 2010;8(1):37–42.

[60] Price MJ, Nayak KR, Barker CM, Kandzari DE, Teirstein PS. Predictors of heightened platelet reactivity despite dual-antiplatelet therapy in patients undergoing percutaneous coronary intervention. Am J Cardiol 2009;103(10):1339–43.

[61] Silvain J, Cayla G, Hulot JS, Finzi J, Kerneis M, O’Connor SA, et al. High on-thienopyridine platelet reactivity in elderly coronary patients: the SENIOR-PLATELET study. Eur Heart J 2012;33(10):1241–9.

[62] Angiolillo DJ, Suryadevara S. Aspirin and clopidogrel: efficacy and resistance in diabetes mellitus. Best Pract Res Clin Endocrinol Metab 2009;23(3):375–88.

[63] Marcucci R, Gori AM, Paniccia R, Giglioli C, Buonamici P, Antoniucci D, et al. Residual platelet reactivity is associated with clinical and laboratory characteristics in patients with ischemic heart disease undergoing PCI on dual antiplatelet therapy. Atherosclerosis 2007;195(1):e217–23.

[64] Gori AM, Cesari F, Marcucci R, Giusti B, Paniccia R, Antonucci E. The balance between pro- and anti-inflammatory cytokines is

associated with platelet aggregability in acute coronary syndrome patients. Atherosclerosis 2009;202(1):255–62.

[65] Antonino MJ, Mahla E, Bliden KP, Tantry US, Gurbel PA. Effect of long-term clopidogrel treatment on platelet function and inflammation in patients undergoing coronary arterial stenting. Am J Cardiol 2009;103(11):1546–50.

[66] Guthikonda S, Lev EI, Patel R, DeLao T, Bergeron AL, Dong JF, et al. Reticulated platelets and uninhibited COX-1 and COX-2 decrease the antiplatelet effects of aspirin. J Thromb Haemost 2007;5(3):490–6.

[67] Cesari F, Marcucci R, Caporale R, Paniccia R, Romano E, Gensini GF, et al. Relationship between high platelet turnover and platelet function in high-risk patients with coronary artery disease on dual antiplatelet therapy. Thromb Haemost 2008;99(5):930–5.

[68] Cesari F, Marcucci R, Gori AM, Caporale R, Fanelli A, Casola G, et al. Reticulated platelets predict cardiovascular death in acute coronary syndrome patients. Insights from the AMI-Florence 2 Study. Thromb Haemost 2013;109(5):846–53.

[69] Marcucci R, Cesari F, Cinotti S, Paniccia R, Gensini GF, Abbate R, et al. ADAMTS-13 activity in the presence of elevated von Willebrand factor levels as a novel mechanism of residual platelet reactivity in high risk coronary patients on antiplatelet treatment. Thromb Res 2008;123(1):130–6.

[70] Mannini L, Marcucci R, Paniccia R, Antonucci E, Giglioli C, Valente S, et al. Erythrocyte deformability and white blood cell count are associated with aspirin resistance in high-risk vascular patients. Clin Hemorheol Microcirc 2006;35(1–2):175–81.

[71] Cecchi E, Marcucci R, Paniccia R, Bandinelli B, Valente S, Giglioli C, et al. Effect of blood hematocrit and erythrocyte deformability on adenosine 5′-diphosphate platelet reactivity in patients with acute coronary syndromes on dual antiplatelet therapy. Am J Cardiol 2009;104(6):764–8.

[72] Homoncik M, Jilma B, Eichelberger B, Panzer S. Inhibitory activity of aspirin on von Willebrand factor-induced platelet aggregation. Thromb Res 2000;99(5):461–6.

[73] Fatini C, Sticchi E, Bolli P, Marcucci R, Giusti B, Paniccia R, et al. Platelet aggregability is modulated by eNOS locus in non-type 2 diabetic patients with acute coronary syndrome. Nutr Metab Cardiovasc Dis 2011;21(1):11–17.

[74] Giusti B, Gori AM, Marcucci R, Saracini C, Vestrini A, Abbate R. Drug and medical device interactions: stent thrombosis and personalizing clopidogrel therapy. Curr Pharmacogenomics Person Med 2010;8(2):124–38.

[75] Giusti B, Gori AM, Marcucci R, Saracini C, Vestrini A, Abbate R. The genetics of antiplatelet therapies. Acute Coron Syndr 2011;10(2):47–54.

[76] Giusti B, Gori AM, Marcucci R, Abbate R. Current status of clopidogrel pharmacogenomics. Pharmacogenomics 2012;13(15):1671–4.

[77] Scott SA, Sangkuhl K, Stein CM, Hulot JS, Mega JL, Roden DM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. Clin Pharmacol Ther 2013;94(3):317–23.

[78] Krishna V, Diamond GA, Kaul S. Do platelet function testing and genotyping improve outcome in patients treated with antithrombotic agents? The role of platelet reactivity and genotype testing in the prevention of atherothrombotic cardiovascular events remains unproven. Circulation 2012;125(10):1288–303.

[79] Gurbel PA, Tantry US. Do platelet function testing and genotyping improve outcome in patients treated with antithrombotic agents? Platelet function testing and genotyping improve outcome in patients treated with antithrombotic agents. Circulation 2012;125(10):1276–87.

[80] Xie HG, Kim RB, Wood AJ, Stein CM. Molecular basis of ethnic differences in drug disposition and response. Annu Rev Pharmacol Toxicol 2001;41:815–50.

[81] Sibbing D, Koch W, Gebhard D, Schuster T, Braun S, Stegherr J, et al. Cytochrome 2C19*17 allelic variant, platelet aggregation, bleeding events, and stent thrombosis in clopidogrel-treated patients with coronary stent placement. Circulation 2010;121(4):512–8.

[82] Richardson TH, Jung F, Griffin KJ, Wester M, Raucy JL, Kemper B, et al. A universal approach to the expression of human and rabbit cytochrome P450s of the 2C subfamily in Escherichia coli. Arch Biochem Biophys 1995;323(1):87–96.

[83] de Morais SM, Wilkinson GR, Blaisdell J, Nakamura K, Meyer UA, Goldstein JA. The major genetic defect responsible for the polymorphism of S-mephenytoin metabolism in humans. J Biol Chem 1994;269(22):15419–22.

[84] Fukushima-Uesaka H, Saito Y, Maekawa K, Ozawa S, Hasegawa R, Kajio H, et al. Genetic variations and haplotypes of CYP2C19 in a Japanese population. Drug Metab Pharmacokinet 2005;20(4): 300–7.

[85] Ferguson RJ, De Morais SM, Benhamou S, Bouchardy C, Blaisdell J, Ibeanu G, et al. A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. J Pharmacol Exp Ther 1998;284(1):356–61.

[86] Xiao ZS, Goldstein JA, Xie HG, Blaisdell J, Wang W, Jiang CH, et al. Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare CYP2C19 mutant allele. J Pharmacol Exp Ther 1997;281(1):604–9.

[87] Ibeanu GC, Blaisdell J, Ghanayem BI, Beyeler C, Benhamou S, Bouchardy C, et al. An additional defective allele, CYP2C19*5, contributes to the S-mephenytoin poor metabolizer phenotype in Caucasians. Pharmacogenetics 1998;8(2):129–35.

[88] Ibeanu GC, Goldstein JA, Meyer U, Benhamou S, Bouchardy C, Dayer P, et al. Identification of new human CYP2C19 alleles (CYP2C19*6 and CYP2C19*2B) in a Caucasian poor metabolizer of mephenytoin. J Pharmacol Exp Ther 1998;286(3):1490–5.

[89] Ibeanu GC, Blaisdell J, Ferguson RJ, Ghanayem BI, Brosen K, Benhamou S, et al. A novel transversion in the intron 5 donor splice junction of CYP2C19 and a sequence polymorphism in exon 3 contribute to the poor metabolizer phenotype for the anticonvulsant drug S-mephenytoin. J Pharmacol Exp Ther 1999;290(2):635–40.

[90] Blaisdell J, Mohrenweiser H, Jackson J, Ferguson S, Coulter S, Chanas B. Identification and functional characterization of new potentially defective alleles of human CYP2C19. Pharmacogenetics 2002;12(9):703–11.

[91] Sim SC, Risinger C, Dahl ML, Aklillu E, Christensen M, Bertilsson L, et al. A common novel CYP2C19 gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. Clin Pharmacol Ther 2006;79(1):103–13.

[92] Rudberg I, Mohebi B, Hermann M, Refsum H, Molden E. Impact of the ultrarapid CYP2C19*17 allele on serum concentration of escitalopram in psychiatric patients. Clin Pharmacol Ther 2008;83(2):322–7.

[93] Hulot JS, Bura A, Villard E, Azizi M, Remones V, Goyenvalle C, et al. Cytochrome P450 2C19 loss-of-function polymorphism is a

major determinant of clopidogrel responsiveness in healthy subjects. Blood 2006;108(7):2244–7.

[94] Fontana P, Hulot JS, De Moerloose P, Gaussem P. Influence of CYP2C19 and CYP3A4 gene polymorphisms on clopidogrel responsiveness in healthy subjects. J Thromb Haemost 2007;5(10):2153–5.

[95] Brandt JT, Close SL, Iturria SJ, Payne CD, Farid NA, Ernest II CS, et al. Common polymorphisms of CYP2C19 and CYP2C9 affect the pharmacokinetic and pharmacodynamic response to clopidogrel but not prasugrel. J Thromb Haemost 2007;5(12):2429–36.

[96] Mega JL, Close SL, Wiviott SD, Shen L, Hockett RD, Brandt JT, et al. Cytochrome p-450 polymorphisms and response to clopidogrel. N Engl J Med 2009;360(4):354–62.

[97] Shuldiner AR, O'Connell JR, Bliden KP, Gandhi A, Ryan K, Horenstein RB, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. JAMA 2009;302(8):849–57.

[98] Kim KA, Park PW, Hong SJ, Park JY. The effect of CYP2C19 polymorphism on the pharmacokinetics and pharmacodynamics of clopidogrel: a possible mechanism for clopidogrel resistance. Clin Pharmacol Ther 2008;84(2):236–42.

[99] Umemura K, Furuta T, Kondo K. The common gene variants of CYP2C19 affect pharmacokinetics and pharmacodynamics in an active metabolite of clopidogrel in healthy subjects. J Thromb Haemost 2008;6(8):1439–41.

[100] Giusti B, Gori AM, Marcucci R, Saracini C, Sestini I, Paniccia R, et al. Cytochrome P450 2C19 loss-of-function polymorphism, but not CYP3A4 IVS10 + 12G/A and P2Y12 T744C polymorphisms, is associated with response variability to dual antiplatelet treatment in high-risk vascular patients. Pharmacogenet Genomics 2007;17(12):1057–64.

[101] Frere C, Cuisset T, Morange PE, Quilici J, Camoin-Jau L, Saut N, et al. Effect of cytochrome p450 polymorphisms on platelet reactivity after treatment with clopidogrel in acute coronary syndrome. Am J Cardiol 2008;101(8):1088–93.

[102] Trenk D, Hochholzer W, Fromm MF, Chialda LE, Pahl A, Valina CM, et al. Cytochrome P450 2C19 681G>A polymorphism and high on-clopidogrel platelet reactivity associated with adverse 1-year clinical outcome of elective percutaneous coronary intervention with drug-eluting or bare-metal stents. J Am Coll Cardiol 2008;51(20):1925–34.

[103] Geisler T, Schaeffeler E, Dippon J, Winter S, Buse V, Bischofs C, et al. CYP2C19 and nongenetic factors predict poor responsiveness to clopidogrel loading dose after coronary stent implantation. Pharmacogenomics 2008;9(9):1251–9.

[104] Gladding P, Webster M, Zeng I, Farrell H, Stewart J, Ruygrok P, et al. The pharmacogenetics and pharmacodynamics of clopidogrel response: an analysis from the PRINC (Plavix Response in Coronary Intervention) trial. JACC Cardiovasc Interv 2008;1(6):620–7.

[105] Jinnai T, Horiuchi H, Makiyama T, Tazaki J, Tada T, Akao M, et al. Impact of CYP2C19 polymorphisms on the antiplatelet effect of clopidogrel in an actual clinical setting in Japan. Circ J 2009;73(8):1498–503.

[106] Price MJ, Murray SS, Angiolillo DJ, Lillie E, Smith EN, Tisch RL, et al. GIFT Investigators. Influence of genetic polymorphisms on the effect of high- and standard-dose clopidogrel after percutaneous coronary intervention: the GIFT (Genotype Information and Functional Testing) study. J Am Coll Cardiol 2012;59(22):1928–37.

[107] Simon T, Verstuyft C, Mary-Krause M, Quteineh L, Drouet E, Méneveau N, et al. French Registry of Acute ST-Elevation and Non-ST-Elevation Myocardial Infarction (FAST-MI) Investigators. Genetic determinants of response to clopidogrel and cardiovascular events. N Engl J Med 2009;360(4):363–75.

[108] Collet JP, Hulot JS, Pena A, Villard E, Esteve JB, Silvain J, et al. Cytochrome P450 2C19 polymorphism in young patients treated with clopidogrel after myocardial infarction: a cohort study. Lancet 2009;373(9660):309–17.

[109] Giusti B, Gori AM, Marcucci R, Saracini C, Sestini I, Paniccia R, et al. Relation of cytochrome P450 2C19 loss-of-function polymorphism to occurrence of drug-eluting coronary stent thrombosis. Am J Cardiol 2009;103(6):806–11.

[110] Sibbing D, Stegherr J, Latz W, Koch W, Mehilli J, Dörrler K, et al. Cytochrome P450 2C19 loss-of-function polymorphism and stent thrombosis following percutaneous coronary intervention. Eur Heart J 2009;30(8):916–22.

[111] Paré G, Mehta SR, Yusuf S, Anand SS, Connolly SJ, Hirsh J, et al. Effects of CYP2C19 genotype on outcomes of clopidogrel treatment. N Engl J Med 2010;363(18):1704–14.

[112] Sofi F, Giusti B, Marcucci R, Gori AM, Abbate R, Gensini GF. Cytochrome P450 2C19*2 polymorphism and cardiovascular recurrences in patients taking clopidogrel: a meta-analysis. Pharmacogenomics J 2011;11(3):199–206.

[113] Mega JL, Simon T, Collet JP, Anderson JL, Antman EM, Bliden K, et al. Reduced-function CYP2C19 genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI: a meta-analysis. JAMA 2010;304(16):1821–30.

[114] Holmes MV, Perel P, Shah T, Hingorani AD, Casas JP. CYP2C19 genotype, clopidogrel metabolism, platelet function, and cardiovascular events: a systematic review and meta-analysis. JAMA 2011;306(24):2704–14.

[115] Mega JL, Close SL, Wiviott SD, Shen L, Walker JR, Simon T, et al. Genetic variants in ABCB1 and CYP2C19 and cardiovascular outcomes after treatment with clopidogrel and prasugrel in the TRITON-TIMI 38 trial: a pharmacogenetic analysis. Lancet 2010;376(9749):1312–9.

[116] Wallentin L, James S, Storey RF, Armstrong M, Barratt BJ, Horrow J, et al. PLATO investigators. Effect of CYP2C19 and ABCB1 single nucleotide polymorphisms on outcomes of treatment with ticagrelor versus clopidogrel for acute coronary syndromes: a genetic substudy of the PLATO trial. Lancet 2010;376(9749):1320–8.

[117] Bouman HJ, Schömig E, van Werkum JW, Velder J, Hackeng CM, Hirschhäuser C, et al. Paraoxonase-1 is a major determinant of clopidogrel efficacy. Nat Med 2011;17(1):110–6.

[118] Cayla G, Hulot JS, O'Connor SA, Pathak A, Scott SA, Gruel Y, et al. Clinical, angiographic, and genetic factors associated with early coronary stent thrombosis. JAMA 2011;306(16):1765–74.

[119] Lewis JP, Ryan K, O'Connell JR, Horenstein RB, Damcott CM, Gibson Q, et al. Genetic variation in PEAR1 is associated with platelet aggregation and cardiovascular outcomes. Circ Cardiovasc Genet 2013;6(2):184–92.

[120] Simon T, Bhatt DL, Bergougnan L, Farenc C, Pearson K, Perrin L, et al. Genetic polymorphisms and the impact of a higher clopidogrel dose regimen on active metabolite exposure and antiplatelet response in healthy subjects. Clin Pharmacol Ther 2011;90(2):287–95.

[121] Gladding P, White H, Voss J, Ormiston J, Stewart J, Ruygrok P, et al. Pharmacogenetic testing for clopidogrel using the rapid

INFINITI analyzer: a dose-escalation study. JACC Cardiovasc Interv 2009;2(11):1095–101.

[122] Barker CM, Murray SS, Teirstein PS, Kandzari DE, Topol EJ, Price MJ. Pilot study of the antiplatelet effect of increased clopidogrel maintenance dosing and its relationship to CYP2C19 genotype in patients with high on-treatment reactivity. JACC Cardiovasc Interv 2010;3(10):1001–7.

[123] Bonello L, Armero S, Ait Mokhtar O, Mancini J, Aldebert P, Saut N, et al. Clopidogrel loading dose adjustment according to platelet reactivity monitoring in patients carrying the 2C19*2 loss of function polymorphism. J Am Coll Cardiol 2010;56(20):1630–6.

[124] Cuisset T, Quilici J, Cohen W, Fourcade L, Saut N, Pankert M, et al. Usefulness of high clopidogrel maintenance dose according to CYP2C19 genotypes in clopidogrel low responders undergoing coronary stenting for non ST elevation acute coronary syndrome. Am J Cardiol 2011;108(6):760–5.

[125] Alexopoulos D, Panagiotou A, Xanthopoulou I, Komninakis D, Kassimis G, Davlouros P, et al. Antiplatelet effects of prasugrel vs. double clopidogrel in patients on hemodialysis and with high on-treatment platelet reactivity. J Thromb Haemost 2011;9(12):2379–85.

[126] Collet JP, Hulot JS, Anzaha G, Pena A, Chastre T, Caron C, et al. CLOVIS-2 Investigators. High doses of clopidogrel to overcome genetic resistance: the randomized crossover CLOVIS-2 (Clopidogrel and Response Variability Investigation Study 2). JACC Cardiovasc Interv 2011;4(4):392–402.

[127] Mega JL, Hochholzer W, Frelinger III AL, Kluk MJ, Angiolillo DJ, Kereiakes DJ, et al. Dosing clopidogrel based on CYP2C19 genotype and the effect on platelet reactivity in patients with stable cardiovascular disease. JAMA 2011;306(20):2221–8.

[128] Collet JP, Cuisset T, Rangé G, Cayla G, Elhadad S, Pouillot C, et al. ARCTIC Investigators. Bedside monitoring to adjust antiplatelet therapy for coronary stenting. N Engl J Med 2012;367(22):2100–9.

[129] Saracini C, Galora S, Maria Palombella A, Abbate R, Giusti B. Pharmacogenomic testing for clopidogrel: state of the ART. Current Pharmacogenomics and Personalized Medicine (formerly Current Pharmacogenomics) 2013;11(3):203–15.

[130] Buchan BW, Peterson JF, Cogbill CH, Anderson DK, Ledford JS, White MN, et al. Evaluation of a microarray-based genotyping assay for the rapid detection of cytochrome P450 2C19 *2 and *3 polymorphisms from whole blood using nanoparticle probes. Am J Clin Pathol 2011;136(4):604–8.

[131] Saracini C, Vestrini A, Galora S, Armillis A, Abbate R, Giusti B. Pharmacogenetics of clopidogrel: comparison between a standard and a rapid genetic testing. Genet Test Mol Biomarkers 2012;16(6):500–3.

[132] Roberts JD, Wells GA, Le May MR, Labinaz M, Glover C, Froeschl M, et al. Point-of-care genetic testing for personalisation of antiplatelet treatment (RAPID GENE): a prospective, randomised, proof-of-concept trial. Lancet 2012;379(9827):1705–11.

[133] Marcucci R, Giusti B, Paniccia R, Giusti B, Gori AM, Valente S, et al. High on-treatment platelet reactivity by ADP and increased risk of MACE in good clopidogrel metabolizers. Platelets 2012;23(8):586–93.

[134] Frelinger III AL, Bhatt DL, Lee RD, Mulford DJ, Wu J, Nudurupati S, et al. Clopidogrel pharmacokinetics and pharmacodynamics vary widely despite exclusion or control of polymorphisms (CYP2C19, ABCB1, PON1), noncompliance, diet, smoking, co-medications (including proton pump inhibitors), and pre-existent variability in platelet function. J Am Coll Cardiol 2013;61(8):872–9.

This page is intentionally left blank

Chapter 26

Pharmacogenetics of Antidepressant Drugs

Chiara Fabbri,[a] Alessandro Minarini,[a] Yoshihiko Matsumoto,[b] and Alessandro Serretti[a]

[a]*Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy,* [b]*Department of Psychiatry, School of Medicine, Yamagata University, Japan*

Chapter Outline

26.1 Overview 543
26.2 Candidate Gene Studies 544
 26.2.1 Antidepressant Pharmacodynamics 544
 26.2.2 Antidepressant Pharmacokinetics 552
26.3 Genome-Wide Approaches 553
 26.3.1 Genome-Wide Association Studies 553
 26.3.2 Pathway Analysis 554
26.4 Conclusion 555
Key Concepts 556
Glossary 556
Acronyms and Abbreviations 556
References 557

26.1 OVERVIEW

At present choosing the antidepressant (AD) to prescribe that is the most appropriate for the individual patient is based only on some general clinical principles, such as the presence of psychiatric comorbidities, depression subtype (e.g., melancholic or anxious depression), the personal or family history of response to a particular drug, or comorbidity with medical diseases. This entails a trial-and-error approach to choosing the AD and the titration regimen, and partly explains the unsatisfactory remission rates of major depressive disorder (MDD) [1]. It is well established that AD efficacy and side effects show a high variability among individuals [2]. Accordingly, standard dose titration may lead to detrimental undertreatment or disabling side effects, and both may result in early treatment discontinuation. For this reason, predictors of treatment outcomes should be identified to aid the choice of tailored treatments. Clinical predictors of AD response have been investigated, but have produced controversial findings, and they cannot be translated into clinical practice [3]. Indeed, MDD phenotypes are very heterogeneous, and clinical features alone are not sufficiently discriminative.

Current evidence suggests that genetic factors (i.e., genetic polymorphisms) contribute substantially to the variance in AD response (≈50% of variance), with mechanisms that are probably similar to those linked to MDD pathophysiology [4,5].

Consistent evidence demonstrates that MDD has a strong genetic background. Indeed, studies in monozygotic and dizygotic twins estimated the hereditability of MDD to be at least 37% [6]. Further, AD response shows a familiar clustering [4], and the metabolism and distribution of several ADs in unaffected twins suggest a substantial genetic component [7]. For these reasons, the effect of genetic polymorphisms on AD efficacy and side effects (i.e., pharmacogenetics) has been a hot topic for the last two decades.

Genetic difference among individuals is determined by di-, tri- and tetra-nucleotide repeats (satellite sequences); large variants >1 kbp due to deletions, insertions or duplications (copy number variants, or CNVs); and nucleotide substitutions. In particular, there are 7 million substitutions in each individual human genome, and over 80% of them take the form of single-nucleotide polymorphisms (SNPs). Thus, SNPs account for over 80% of the variability among humans, including disease predisposition [8]. As of this writing, the HapMap project has mapped the location of over 3 million SNPs (http://hapmap.ncbi.nlm.nih.gov/). Given their pivotal contribution to unique genetic make-up, SNPs have been the most investigated genetic variant as predictors of AD clinical outcomes.

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00026-8
© 2014 Elsevier Inc. All rights reserved.

The first pharmacogenetic studies of AD response were based on the candidate gene approach—that is, the analysis of genetic variants selected a priori on the basis of molecular and pharmacological data and/or animal models. Genes selected this way include those coding for key molecular components of the serotoninergic, noradrenergic, and dopaminergic systems, functional abnormalities in which are thought to contribute to MDD pathophysiology (monoaminergic theory). Their restoration is one of the consequences of AD treatment. According to monoaminergic theory, MDD is caused by decreased monoaminergic neurotransmission in the brain (although this is without doubt an oversimplification) [5]. Thus, the serotonin transporter gene (*SLC6A4*), serotonin 1A receptor (*HTR1A*), serotonin 2A receptor (*HTR2A*), tryptophan hydroxylase 1 and 2 (*TPH1*, *TPH2*), catechol-O-methyltransferase (*COMT*), monoamine oxidase A (*MAOA*), and norepinephrine transporter (*SLC6A2*) were considered among the best a priori candidates. Nevertheless, given the inconsistency of findings obtained by the candidate gene approach, it became clear that the genetic component of AD response is likely linked to a number of genes, each with a small effect size, rather than to one or few major loci.

Recent pharmacogenetic studies have extended the investigation of genetic variability to the whole genome, through genome-wide association studies (GWASs), which allow the genotyping of hundreds of thousands of polymorphisms across the whole genome through microarray technology. The main advantage of GWASs is the high number of loci genotyped without the need for a priori hypotheses. This is a very useful tool because AD mechanisms of action are not fully understood. Unfortunately, GWAS results have fallen below expectations, probably because of some unsolved methodological and technical issues (see Section 26.3.1). Strategies to solve these issues are discussed in Section 26.4.

Despite the inconsistency of previous results and the reported need for methodological improvements, candidate gene studies and GWASs provided well-supported evidence of involvement of genes relating to the monoaminergic system, neurotrophic system, hypothalamus–pituitary–adrenal (HPA) axis, and drug transport/metabolism in AD efficacy (see Tables 26.1 and 26.2 for summaries). The following sections provide a critical review of the main findings in the field of AD pharmacogenetics and discuss the challenge of translating genetic information into clinical applications—the prescribing of gene-guided tailored therapies.

TABLE 26.1 Summary of Candidate-Gene Study Results

Gene	
Polymorphisms Studies with More Than 50% Positive Results	
SLC6A4	5-HTTLPR
COMT	rs4680
IL1B	rs16944
BDNF	rs6265
GNB3	rs5443
Polymorphisms Studies with Equal Number ±1 of Positive and Negative Results	
SLC6A4	STin2 VNTR
TPH1	rs1800532
HTR1A	rs6295;rs1800042
SLC6A2	rs5569;rs2242446
CRHR1	rs242941(rs1876828 and rs242939 in LD)
FKBP5	rs1360780(rs4713916 and rs800373 in LD)
DTNBP1	rs760761; rs2619522
ABCB1	rs2032583;rs2235040
Polymorphisms Studies with More than 50% of Negative Results	
SLC6A4	5-HTTLPR/rs25531 (triallelic locus)
HTR2A	rs7997012;rs6311;rs6313
HTR6	rs1805054
MAOA	VNTR 1.2 kb upstream the coding sequence
GRIK4	rs1954787
ABCB1	rs2032582

Note: Results shown are according to their degree of replication. Only polymorphisms investigated by at least three studies were reported.

26.2 CANDIDATE GENE STUDIES

26.2.1 Antidepressant Pharmacodynamics

In the following subsections, genes involved in antidepressant pharmacodynamics are discussed. Pharmacodynamics includes all biological processes that are related to antidepressant mechanisms of action; that is, molecular targets that are directly modulated by antidepressant drugs (e.g., cell receptors) and the subsequent molecular events that lead to the pharmacological effects.

26.2.1.1 *Monoaminergic System*

According to monoaminegic theory, MDD develops because of a deficit in noradrenergic, dopaminergic, and/or serotonergic neurotransmission. The majority of currently available ADs have been developed on the basis of

TABLE 26.2 Summary of GWAS Results

Study	Top Genes	Phenotype	Biological Rational for Association
STAR*D	*UBE3C*	AD efficacy	Modifies proteins to signal them for degradation; no clear biological link with phenotype
	BMP7	AD efficacy	Induces endochondral osteogenesis; no clear biological link with phenotype
	RORA	AD efficacy	Modulation exerted by ADs may restore circadian rhythm abnormalities of MDD
	ACSS3	AD efficacy	Involved in lipid synthesis and energy generation; no clear biological link with phenotype
	NEDD4L x quality of life	AD efficacy	Modifies proteins to signal them for degradation; no clear biological link with phenotype
	SACM1L	Sexual side effects	Regulates vesicle trafficking in endoplasmic reticulum and Golgi apparatus; no clear biological link with phenotype
	EMID2	Vision/hearing side effects	Involved in the regulation of corneal collagen fibrillogenesis
MARS	*CDH17*	AD efficacy	Calcium-dependent cell adhesion protein; may have an important role in synapse formation in the hippocampus
	EPHB1	AD efficacy	One of the ephrin receptors; regulates chemotaxis, proliferation, and polarity of hippocampus neural progenitors
GENDEP	*UST*	AD efficacy	Involved in synthesis of extracellular matrix components; may be linked to *CSGALNACT1* (see *CSGALNACT1* function)
	IL11	AD efficacy	Involved in inflammation processes, in line with inflammatory theory of MDD
GENDEP + STAR*D	*CSGALNACT1*	AD efficacy	Involved in chondroitin chain biosynthesis; may play some role in neuroplasticity, regeneration, and brain development

Note: Genes reported as "top genes" do not necessarily represent genome-wide significant results. AD—antidepressant; MDD—major depressive disorder.

this assumption, despite the fact that it fails to completely elucidate the disease pathophysiology.

Given the key role of serotonin (5-HT) in the regulation of mood and emotions, [9] genes relating to the serotoninergic system were heavily investigated, among them the serotonin transporter gene (*SLC6A4*), which was hypothesized as an optimal candidate. The serotonin transporter (SERT) regulates brain 5-HT neurotransmission by transporting

5-HT from the synaptic cleft to presynaptic neurons. The blocking of SERT, the main target of ADs, can lead to recovery thanks to an increase in 5-HT availability at the synaptic level.

The most investigated variant in *SLC6A4* is a 44-bp insertion/deletion polymorphism (5-HTTLPR) involving 2 units in a sequence of 16 repeated elements in the promoter region. 5-HTTLPR has a functional effect: the 16-repeat sequence is called the long allele (L) and shows a twice-basal SERT expression compared to the 14-repeat sequence, called the short allele (S) [10]. Interestingly, imaging studies show that S carriers have higher levels of amygdala activity and reduced functional coupling of the amygdala–anterior cingulate cortex (ACC) circuit [11]—phenotypes associated with anxiety and MDD. It was hypothesized that S carriers may be less able to respond to repeated releasing inputs of 5-HT because of the low activity of SERT; the consequent reduction of serotoninergic transmission may be linked to an increased risk of MDD.

For the most part, pharmacogenetic studies demonstrated a poorer AD response in SS carriers, particularly in Caucasians treated with SSRIs, compared to L carriers [12]. A possible explanation is that L carriers have a more dynamic serotonergic system—that is, higher SERT functionality may allow a better balancing of the serotonergic tone (in terms of time and amount) in response to endogenous and exogenous stimuli. The different effects across populations may be due to wide variations in 5-HTTLPR allele frequencies among ethnic groups (the S allele is present in 42% of Caucasians but 79% of Asians) [13].

As a key regulator of SERT activity, 5-HTTLPR may also influence the risk of AD-induced side effects, particularly in the case of SSRIs. Most available studies suggested that the S allele is a risk factor; however, some reported no association (although the latest group of studies was mainly performed on non-Caucasian samples) [14]. This further suggests a population-specific effect of the polymorphism.

Despite the interesting role of 5-HTTLPR, polymorphisms are often not independent; moreover, the SNP rs25531 A/G was reported to lie in the 5-HTTLPR sequence, with a possible influence on the functional effect of 5-HTTLPR itself [15]. Indeed, the rs25531 G variant in association with the L allele (LG) may result in reduced expression of *SLC6A4*, equivalent to that conferred by the S allele [15]. Moreover, single-photon emission-computed tomography imaging suggested that LA/LA carriers may have a more dynamic serotonergic system that seems to confer a higher probability of SSRI response [16]. However, this hypothesis was not supported by pharmacogenetic studies, which showed mainly negative results [12].

Further studies are required to clarify the relative contribution of and interaction among *SLC6A4* polymorphisms; the finding that rs25531 may lie 18 bp 5′ from the site of the L/S repeat and not in it [17] calls into question previous findings.

Another variant of *SLC6A4* influencing SERT expression is a 17-bp VNTR (variable number of tandem repeats) in intron 2 (STin2), which comprises 9, 10, or 12 copies of a 16–17-bp repeat and may influence gene transcription synergistically with 5-HTTLPR [18]. In particular, the 12-repeat variant was demonstrated to cause higher gene expression *in vitro* [19] and *in vivo* [20]. Investigation of the association between this polymorphism and AD response provided mainly negative results, while positive findings suggested a better response in long-allele carriers among Asians and in short-allele carriers among Caucasians [12]. Meta-analytic results further suggested an opposite effect of the variant in Caucasians versus Asians, but results were not unequivocal [21]. Finally, no effect of STin2 on AD-induced side effects was demonstrated [21].

5-HT receptors represent other pivotal regulators of 5-HT neurotransmission, and the association of their variants and AD response was thoroughly investigated.

The serotonin 1A receptor (5-HT1A receptor), encoded by the *HTR1A* gene, is an inhibitory autoreceptor widely distributed in regions that receive serotonergic input from the raphe nuclei: frontal cortex, septum, amygdala, hippocampus, and hypothalamus, and it can be expressed both pre- and postsynaptically. This receptor modulates the serotonergic activity of these areas by reducing the firing rate of raphe nuclei neurons, the amount of 5-HT released per action potential, and the synthesis of the neurotransmitter. The involvement of this gene in AD response is supported by evidence showing that several ADs desensitize raphe 5-HT1A autoreceptors, leading to an enhancement of serotonergic neurotransmission that may be responsible, at least partially, for the AD effect. Furthermore, the blocking of 5-HT1A autoreceptors may accelerate AD action [22].

The most investigated *HTR1A* variants are rs6295 (1019C/G) and rs1800042 (Gly272Asp). The rs6295 G allele results in upregulation of the gene; thus, it may inhibit the therapeutic effect of ADs. Several studies consistently suggested the G allele as a risk factor for MDD, suicidal behavior, and AD nonresponse [23], although some obtained conflicting results [24–27]. Interestingly, an association between rs6295 and AD response may exist only in particular subgroups of patients, such as females [28] and those with melancholic depression [29], confirming the importance of demographic and clinical modulators in pharmacogenetic studies. Regarding rs1800042, initial suggestions that it may modulate AD response [30] were not confirmed [28,31].

Another relevant player in the serotonin synapses is the 5-HT2A receptor (*HTR2A* gene), one of the main postsynaptic serotonin receptors that is distributed widely throughout the CNS with high densities in the frontal cortex. An increasing amount of evidence suggests that 5-HTR2A is involved in MDD pathophysiology and AD mechanisms of action. Indeed, decreased cortical 5-HT2A receptor binding in MDD patients compared to healthy controls was reported [32]. With regard to *HTR2A* polymorphisms, the A allele of the intronic rs7997012 showed a favorable effect in a large MDD sample treated with SSRIs (STAR*D; formally, sequenced treatment alternatives to relieve depression) according to two studies based on independent genotyping assays [33,34].

On the other hand, in samples treated with non-SSRIs or mixed ADs, negative [25,35–38] or opposite [39] results were obtained. Despite negative findings, other SNPs in *HTR2A* (rs9316233 and rs2224721) were implicated in AD efficacy [36], suggesting the gene's contribution to SSRI response through a multilocus model. In other words, inconsistent results on rs7997012 may be partly due to complex interactions with other *HTR2A* SNPs, other genes, or even nongenetic confounders. Indeed, the effect of rs7997012 on AD response may be influenced by an interaction with *HTR2A* rs6311 (1438G/A) and gender [40]. In turn, the effect of rs6311 on AD response may be modified by other polymorphisms. First, rs6311 may impact treatment outcome through an interaction with *GNB3* (guanine nucleotide binding protein [G protein], beta polypeptide 3) rs5443 (C825T) and *SLC6A4* rs25533 [41]. Second, rs6311 is in linkage disequilibrium (LD) with rs6313 (102T/C) [42], and the effect of rs6313 on AD response may depend on the *GRIK4* (glutamate receptor, ionotropic, kainate 4) [43]. Consistent with the hypothesis of a multilocus model, studies that focused only on rs6311 and/or rs6313 did not show any clear evidence of association with AD efficacy [12].

Despite negative reports [44–47], the involvement of rs6311/rs6313 in antidepressant-induced side effects is possible. Indeed, the rs6311 GG and rs6313 CC genotypes may be risk factors for side effect development [21]. Regarding other serotoninergic receptors, preliminary findings have been reported primarily for 5-HT3 and 5-HT6.

The 5-HT3 receptor is expressed throughout the central and peripheral nervous systems and mediates a variety of physiological functions. It is the only ion channel receptor among the serotonin receptors. Five subunits (A–E) of the 5-HT3 receptor have been identified. The genes encoding them have been investigated mainly in regard to AD side-effect profiles rather than response. In particular, *HTR3A* rs1062613 (C178T) and two variants in *HTR3B* (−100 − 102 AAG deletion and rs1176744) have been associated with vomiting and nausea caused by paroxetine intake [46,48,49].

In regard to the 5-HT6 receptor, which is encoded by the *HTR6* gene, several studies reported associations between this gene and brain-related features such as cognitive function, personality traits, and mood regulation [50,51]. 5-HT6 appears to be involved in AD mechanisms of action as well. Indeed, its stimulation causes antidepressant-like behavioral and biochemical modifications, suggesting that this receptor contributes to serotonergic modulation of clinically relevant psychopharmacological processes [52]. In agreement with this hypothesis, the rs1805054 variant (C267T) may modulate AD response [53,54], although further investigations are required to clarify the role of this SNP.

So far we have focused on 5-HT neurotransmission, but another key regulation point of 5-HT activity is the biosynthesis of 5-HT itself. The limiting step of 5-HT biosynthesis is catalyzed by tryptophan hydroxylase (TPH), which is codified by two distinct genes, *TPH1* and *TPH2*. *TPH1* is ubiquitous but predominantly expressed in peripheral organs, while *TPH2* is more selectively expressed in the brain. The rationale behind the investigation of these genes as predictors of AD response derived from the hypothesis that a reduced basal 5-HT synthesis may be a risk factor for MDD.

Regarding *TPH1*, available data mainly pertain to rs1800532 (A218C), a biallelic SNP located in a potential GATA transcription-factor binding site. The A allele is associated with decreased 5-HT synthesis and, according to monoaminergic theory, may determine insufficient AD efficacy [55]. This hypothesis was followed by many pharmacogenetic studies, unfortunately with conflicting results [12]. As for AD-induced side effects, a possible effect of rs1800532 on AD-induced weight gain was only weakly supported [56].

Data on *TPH1* suggest the need for further investigation, in particular because of possible interactions between *TPH1* polymorphisms and polymorphisms in related genes, such as *HTR2A*, *GNB3*, and *COMT* [56,57]. The same is true for *TPH2* since several variants were investigated but with conflicting results [12].

The central serotoninergic system is deeply interrelated with the other monoaminergic circuits. Therefore, the genetic variants reported so far likely interact with polymorphisms in key noradrenergic and dopaminergic genes, contributing to the final clinical phenotype.

The noradrenergic system has been widely investigated because it is one of the main targets of tricyclic ADs. Among the key genes in this system, the most investigated are catechol-O-methyltransferase (*COMT*), monoamine oxidase A (*MAOA*) (both clearly involved in other systems such as the dopaminergic system), and norepinephrine transporter (*SLC6A2*) genes, which code for the main enzymes responsible for amine metabolism.

The catechol-O-methyltransferase gene (*COMT*) plays a key role in monoamine metabolism and seems to be involved in MDD pathophysiology and AD response. Two different transcription sites are responsible for the production of the soluble cytoplasmic (S-COMT) and membrane-bound (MB-COMT) isoform. MB-COMT is believed to be the dominant isoform in the brain, whereas S-COMT predominates in the periphery. MB-COMT assumes greater importance in the frontal cortex and in the striatal neurons postsynaptic to the dopaminergic neurons, where it is responsible for more than 60% of dopamine degradation.

The rs4680 (Val108/158Met) variant has been widely studied because it exhibits a relevant functional effect: the Val/Val genotype catabolizes dopamine at up to four times the rate of the Met/Met homozygote, resulting in a significant reduction of synaptic dopamine following neurotransmitter release. This polymorphism may affect not only synaptic dopamine levels but also serotoninergic activity. It has been hypothesized that rs4680 may affect SSRI response through the modulation of dopamine bioavailability in the prefrontal cortex [58]. Available data mainly reported the Met variant as the favorable allele for AD response, with an allele dose effect (better outcome in Met/Met carriers and intermediate outcome in Met/Val carriers) that confirms the previous hypothesis [59]. There are also preliminary findings about rs4680's possible effect on the risk of weight gain during treatment with various ADs [56]. However, recent negative results indicate the need for further studies [27].

Recent improved coverage of *COMT* genetic variability confirmed the hypothesis that this gene may modulate AD efficacy. Other variants were implicated (rs4633, rs4818, rs165599, rs165774, rs174696, and rs13306278) [35,60,61]. Interestingly, rs13306278 was also associated with an altered ability to bind nuclear proteins [61].

Monoamine oxidase (MAO) is an enzyme family with a key role in the inactivation of neurotransmitters and has a direct impact on the regulation of neurotransmitter levels in the brain. In humans, two distinct isoforms are expressed: MAO-A, the most investigated, mainly breaks down serotonin, norepinephrine, and epinephrine; MAO-B, mainly investigated for its role in Parkinson's disease, breaks down phenethylamine and benzylamine. Both forms equally break down dopamine. The role of MAO-A in MDD pathophysiology and AD response is largely supported by the clinical effectiveness of MAO-A inhibitors (MAOIs), on which pharmacogenetics has provided some interesting findings.

One particular polymorphism in the *MAOA* gene that has been widely investigated is a 30-bp VNTR located 1.2 kb upstream in the *MAOA* coding sequence that seems to influence the transcription rate of the gene and has been linked to variations in enzyme activity [62]. Indeed, *in vitro* studies demonstrated that repeat sequences with 3.5 or 4 copies are transcribed 2–10 times more efficiently than those with 3 or 5 copies [63]. Unfortunately, no clear confirmation has been provided *in vivo* [64,65]. Pharmacogenetic studies retrieved mainly negative findings [12], while some data supported insufficient response in longer alleles [66], with a possible selective effect in females [67,68]. The reported sex-specific effect is not surprising considering that the gene is located on the X chromosome and that neurotransmitter turnover is influenced by sex hormones [69].

Norepinephrine (NE) has a primary role in the regulation of energy and alertness [9], and the beta1 adrenergic receptor (*ADRB1* gene) has been the most investigated candidate. *ADRB1* is an important regulator of mood, memory, neuroendocrine activity, and autonomic function; moreover, it is involved in the mediation of AD response [70]. The functional SNP rs1801253 (G1165C) was associated with enhanced coupling to the stimulatory Gs protein and with increased adenylyl cyclase activation, resulting in better and faster AD response [71].

Few data are available on the alpha2 adrenergic receptor (*ADRA2A* gene), which may influence the risk of weight gain during mirtazapine treatment [72] because it acts as an antagonist of central presynaptic alpha2 receptors. Furthermore, *ADRA2A* is expressed in hippocampal progenitor cells and inhibits hippocampal neurogenesis. The blockade of *ADRA2A* was consistently demonstrated to enhance the AD effect of impipramine [73], suggesting the gene as a coregulator of AD effect.

The norepinephrine transporter (NET) is encoded by the *SLC6A2* gene. It is a monoamine transporter that reuptakes NE from the synapse back to cytosol; hence, other transporters, known as VMATs (vesicular monoamine transporters), sequester NE into vesicles for later storage and release. NET represents one of the primary targets of ADs, especially secondary amine tricyclics [74]. In Asians, rs5569 (G1287A) may be a predictor of response, particularly to noradrenergic antidepressants [62,75,76]. A higher number of variants in Caucasians were investigated, but results so far do not allow any conclusion to be reached [58].

The dopaminergic system has received less attention in the field, although both preclinical [77] and clinical data [78] demonstrated its involvement in the pathogenesis of MDD. In detail, an underactivation of the nigro-striatal dopaminergic system was demonstrated in MDD [78]. An increase of dopaminergic neurotransmission after chronic AD treatment [78,79] and the reversal of AD effect by acute administration of a D2 receptor selective antagonist [80] were also demonstrated.

Dopamine receptors are divided into the D1-like family (D1 and D5, which are coupled to a Gs protein and activate adenylate cyclase) and the D2-like family (D2, D3, and D4, which are coupled to a Gi protein and inhibit adenylate cyclase). Only the D2-like family has been associated with depressive disorders. Despite evidence of involvement of the dopaminergic system in MDD and AD response, few pharmacogenetic data are available. The D2 receptor (DRD2 gene), which plays a pivotal role in hedonic behavior, has no pharmacogenetic evidence of association with AD efficacy [81,82], with the exception of a possible effect on early symptom improvement during AD treatment [83].

A role for the D3 receptor in AD effect was suggested by the demonstration of long-lasting changes in the rat brain during D3 agonist administration, similar to what was observed during AD treatment [84]. Data on the *DRD3* gene are few and do not support its involvement in AD response modulation [82]. Finally, the VNTR polymorphism, located at the third exon of the *DRD4* gene, shows a weak association with AD response [85].

The final important player in the dopaminergic system is the dopamine transporter (DAT, coded by the *SLC6A3* gene), which is responsible for dopamine reuptake from the synaptic cleft into the presynaptic neurons. Some evidence supports the involvement of DAT in synaptic plasticity in brain regions that are involved in MDD pathophysiology (e.g., the hippocampus) and in neuronal development. Furthermore, DAT might modulate AD clinical effects. Interestingly, it was reported that AD drugs modulate the availability of DAT, which supports the role of the dopaminergic system in AD mechanisms of action [80,86]. The most investigated polymorphism in the *SLC6A3* gene is a 40-bp VNTR in exon 15, which was reported to affect DAT expression [87]. The 9/10 and 9/9 genotypes may be associated with the risk of poorer and slower response to AD drugs [88].

26.2.1.2 Glutamatergic System

Increasing evidence suggests that the glutamatergic system plays a role in MDD pathogenesis and in the mechanisms of AD response. The vast majority of excitatory neurotransmission in the CNS is mediated by aminoacid glutamate, which acts both on ionotropic receptors (NMDA, AMPA, and kainate receptors) and on receptors linked to intracellular second messenger systems (metabotropic, or mGlu). Glutamatergic theory posits that glutamate may shape the risk of MDD, influencing neuronal fate (neurotoxicity due to excess influx of Ca++ through the NMDA receptor) or the unfolding of new neuronal nets (neuroplasticity). The greatest evidence of imbalance in glutamatergic signaling was demonstrated in the hippocampus, where it is likely involved in stress-induced neuronal atrophy and death in the CA3 layer (the region of the dentate gyrus) [89]. At this level, the glutamatergic system modulates the growth of neurons and dendrites to form neuronal nets that, thanks to long-term potentiation (LTP), translate salient stimuli in permanent or semipermanent informative biological structures.

During depressive states, the chronic activation of the system, as demonstrated by proton magnetic resonance spectroscopy [90], may be responsible for neuronal death and the subsequent decrease in hippocampal volume. Convincing evidence supports the involvement of glutamatergic pathways in AD mechanisms of action. Indeed, substances with glutamatergic properties have AD effects; [91] for example, some AD treatments lead to the downregulation of the glycine B site (glycine is a modulator of glutamatergic function) and NMDA receptors. The antidepressant tianeptine is also a modulator of glutamatergic function [92] and was demonstrated to target the diffusion of AMPA receptors on cell surfaces [93]; finally, compounds that target the glutamatergic system have consistently been shown to have an antidepressant effect.

Ketamine, which acts through inhibiting NMDA receptors, reverses the behavioral and physiological alterations induced by chronic mild stress in rats, and it has been proved to be an antidepressant in humans [94]. The mechanism of the ketamine antidepressant effect has not as yet been clarified, but it was recently proposed that both NMDA antagonism and AMPA receptor activation are involved. The NMDA receptors have a slower and more prolonged postsynaptic current than do the AMPA receptors, which induce a more rapid dissociation of glutamate; the antidepressant effect probably results from the "correct" balance of NMDA and AMPA receptor activity in definite brain areas. In any case, the antidepressant effect of glutamatergic substances may be partially due to monoaminergic mechanisms: downregulation of the adrenergic receptors and enhancement of the serotoninergic function are associated with administration of glutamatergic ADs. On the other hand, chronic treatment with monoaminergic ADs causes a reduction in glutamate release [94].

Despite the relevance of the glutamatergic system in MDD and AD mechanisms of action, there are as yet few studies focused on glutamatergic variants, and those that are available investigated individual polymorphisms. In detail, *GRIK4* (glutamate receptor, ionotropic, kainate 4) was repeatedly investigated for its putative role in susceptibility for depression because of the suggested antidepressant-like phenotype of *GRIK4*(−/−) mice [95]. Unfortunately, pharmacogenetic results are sparse and inconsistent. Indeed, some studies associated different SNPs (rs1954787 and rs12800734) with AD efficacy without replicated results, while others produced negative findings both for *GRIK4* and for all glutamate NMDA and kainate receptor genes [12]. Interestingly, the GRM7 gene (glutamate receptor, metabotropic 7) may be a predictor of rapid antidepressant response [96]. This is consistent with the evidence that glutamatergic ADs (such as ketamine) are faster acting than their monoaminergic counterparts. Other genes (*GRIK2* and *GRIA3* (glutamate receptor, ionotrophic, AMPA 3)) were associated with treatment-emergent suicidal ideation and AD-induced sexual dysfunction (*GRIK2*, *GRIA1*, *GRIA3*, and *GRIN3A* (glutamate receptor, ionotropic, N-methyl-D-aspartate 3A)), but confirmation of these associations is still needed [58].

The genes discussed so far represent only the first step toward understanding the glutamatergic molecular cascade. A number of interrelated genes are still to be investigated. For example, an interesting line of research points toward the *KCNK2* gene (potassium channel subfamily K member 2), which codes for a 2-pore domain potassium channel involved in the regulation of excitability and the resting potentials of neurons. The opening of these channels may be inhibited by stimulation of the metabotropic glutamate receptors mGluR1 and mGluR5. *KCNK2* is expressed in brain areas implicated in MDD, including the prefrontal cortex, the hippocampus, and other limbic structures. Animal models suggest its involvement in the pathophysiology of MDD and AD response: *KCNK2* KO mice show increased efficacy of 5-HT neurotransmission, resistance to depression, and a substantially reduced elevation of corticosterone levels under stress.

In humans, several SNPs in the *KCNK2* gene were found to be predictors of nonremission after first- and second-line treatments, suggesting that the gene plays a role in TRD (treatment-resistant depression) [97]. This result suggests that the glutamatergic system may represent an alternative molecular pathway for restoring the biological alterations of MDD when the monoaminergic system is "poorly responsive." This hypothesis is consistent with the evidence of ketamine efficacy in TRD. It is important that the frequent inconsistencies in findings involving the glutamatergic and monoaminergic systems, when individually investigated, point to the need for an integrated perspective to provide a preferential point of observation for unraveling the molecular events that drive AD effect. Indeed, these neuronal systems are involved in a complex coregulation.

26.2.1.3 HPA Axis and Inflammation

The inflammatory theory of depression posits that an imbalance in the immune mediators is involved in the pathogenesis of depressive symptoms. Indeed, anyone who has ever had the flu knows that being sick brings with it a loss of appetite and loss of interest in the physical and social environment. These symptoms are similar to those of a major depressive episode. The inflammatory theory is supported by the finding of elevated levels of proinflammatory cytokines and other inflammation-related proteins in the plasma and cerebrospinal fluid (CSF) of MDD subjects as well as in the postmortem brain. ADs can lead to normalization of these parameters [5].

A potent activator of proinflammatory cytokines has been consistently found to induce mild fever, anorexia, anxiety, depressed mood, and cognitive impairment in human volunteers [98]. Moreover, the levels of anxiety, depression, and cognitive impairment are associated with levels of circulating cytokines [98,99], and the administration of cytokines (e.g., INF-α, which is used in the treatment of hepatitis C or malignant melanoma) may cause depressive symptoms, including suicidal ideation [5]. Proinflammatory cytokines cause various clinical aspects of depression, including hyperactivity of the HPA, disturbed 5-HT metabolism, and neurovegetative symptoms. It is known that one-third of MDD subjects have elevated peripheral inflammatory biomarkers, even without a medical illness; a dysregulation of the hypothalamic–pituitary–adrenal (HPA) axis (which contributes to the regulation of inflammatory cell activity and cytokine release) was reported in up to 70% of patients with MDD [100]. Finally, MDD treatment outcome is influenced by the AD-induced modulation of cytokines, obtained through direct or indirect action on intracellular cyclic adenosyl monophosphate (cAMP), 5-HT metabolism, and HPA axis functioning [101].

The main neuroendocrine regulator of the HPA axis is the corticotrophin-releasing hormone (CRH); CRH receptors 1 and 2 (*CRHR1* and *CRHR2*) mediate the effect of glucocorticoids in the CNS. *CRHR1* antagonists have AD effect in both animals and humans [102], suggesting a key role for CRH and its receptors in depression. Concerning CRHR1, some polymorphisms (rs242941, rs1876828, and rs242939) were thought to be related to SSRI response [103], especially in a cluster of patients with anxious depression; however, no confirmation of this result was found [104]. A different SNP was later implicated in AD response [105].

Studies on the *CRHR2* gene suggested a possible involvement of rs2270007 [104] and rs2267716–rs255105 [105] in SSRI response, and the possible involvement of rs917195 in TCAs efficacy [106]. Hyperactivity in the HPA axis in MDD might also be caused by impaired glucocorticoid signaling. Glucocorticoids act through the glucocorticoid receptor (GR), which is coded by the *NR3C1* gene. The GR is a ligand-activated transcription factor. Thus, genetic variants occurring in the NR3C1 gene may significantly affect glucocorticoid sensitivity. Several NR3C1 polymorphisms were consistently associated with depression as well as AD response.

In detail, the ER22/23EK polymorphism, which consists of two linked point mutations in codons 22 and 23, was associated with faster clinical response [107]. The BclI polymorphism, which identifies two alleles with fragment lengths of 4.5 and 2.3 kb, gained attention because replicated evidence indicated that it may be linked to altered GR function [108–110]. Unfortunately, negative findings were provided for BclI as well as for other polymorphisms in *NR3C1* [12]. All the data reported up to now refer to candidate gene studies, while a recent GWAS provided new candidate SNPs (rs852977, rs10482633, and rs10052957) [36].

In the absence of hormones, the GR resides in the cytosol complex with a variety of proteins, including *FKBP5* (FK506-binding protein 52), an established regulator of GR sensitivity. Given its key role, the *FKBP5* gene has received even more investigative attention than *NR3C1* as a predictor of AD efficacy. Unfortunately, results were mainly negative even though some candidate SNPs (rs1360780, rs3800373, and rs4713916) emerged [104,111–114]. Preliminary findings also exist for the CRH-binding protein gene that encodes for a plasma protein involved in the inactivation of CRH [105].

It is now clear that the brain is not a sanctuary for the immune system, as hypothesized in the past; however, inflammatory cytokines have been shown to access the brain and can virtually interact with every pathophysiologic domain known to be involved in MDD, including neurotransmitter metabolism, neuroendocrine function, and neural plasticity [115]. The most investigated cytokine gene is interleukin 1 beta (*IL1B*), which shows multiple and reciprocal interactions with the monoaminergic, cholinergic, and GABAergic systems in the CNS [58]. Pharmacogenetic results are quite promising for this gene: the promoter rs16944 was selectively associated with SSRI response [116–118], and the intronic rs1143643 [118] may be a predictor of AD response. rs1126757 is also a good candidate [119].

An interesting regulator of the release of inflammatory cytokines such as *IL1B* is the purinergic receptor P2X, which is a key mediator in chronic inflammation, neurodegeneration, and chronic pain. P2X7 receptor-deficient mice show a substantially attenuated inflammatory response [120]. Although the expression of this nuclear receptor in neurons is still uncertain, its activation can mediate and/or enhance neurotransmitter release (e.g., glutamate) and may affect neuronal plasticity and neuronal cell death. Thus, P2X7 receptors may act as mediators between the immune and the nervous system and play a role in both neurodegenerative and psychiatric disorders. Some functional polymorphisms identified in the *P2RX7* gene seem to increase the risk of a familial mood disorder and time spent ill [121]. On the basis of these data, further investigation is needed to clarify the contribution of *P2RX7* and other inflammatory genes to AD response.

26.2.1.4 Neurotrophic Factors and Signal Transduction

Neurotrophic factors were first characterized as regulators of neural growth and differentiation during development, but today they are also known to be potent regulators of plasticity and survival of neurons and glia during adulthood. A role for neuronal plasticity in affective disorders was hypothesized after the observation that hippocampus atrophy following stress was reversed by ADs in parallel with an increase in the expression of neurotrophic factors, especially brain-derived neurotrophic factor (*BDNF*) [122]. Indeed, ADs can reverse a decrease in volume of specific brain areas in MDD (the hippocampus, left anterior cingulate gyrus, subgenual region, area 9 (dorsolateral cortex), area 10/47 (rostral orbital cortex), and area 47 (caudal orbital cortex)) [5].

BDNF is the gene that has attracted the most research attention. It is a member of the nerve growth factor (NGF) superfamily, which was demonstrated to be underexpressed during depressive states [123]. AD treatments are consistently associated with increased *BDNF* expression in the hippocampus of adult rats [124]. As a consequence of these findings, *BDNF* polymorphisms have been extensively studied, especially rs6265, a valine to methionine (V66M) substitution in position 196. The Met allele was associated with poorer episodic memory and abnormal hippocampal activation [125]. Pharmacogenetic studies mainly suggested a positive molecular heterosis effect, or a more favorable outcome in Met allele carriers [12]. On the other hand, some studies reported a better response in the Val/Val genotype, and several negative findings exist [12]. There are various reasons for these contradictory results.

One reason is the different ethnicities of the examined samples—considerable *BDNF* allele and haplotype diversity among global populations was reported [126]. Furthermore, different ADs may have different effects on central *BDNF* expression, so the heterogeneity in treatments both among and within samples represents another possible confounding factor. Other polymorphisms in the gene were associated with AD response, but findings are still preliminary. Moreover, the *NTRK2* gene (encoding for the *BDNF* receptor) received attention as a modulator of AD response because of its possible interaction with the *BDNF* gene [62].

Vascular endothelial growth factor (*VEGF*) was first characterized (and named) for its role in vascular permeability, and was later described as a potent endothelial cell mitogen and survival factor. Recently, it was found to be influential beyond the vascular system. Indeed, in addition to regulating endothelial cells and blood vessel formation/permeability, *VEGF* has been hypothesized to be a true neurotrophic factor. Indeed, animal studies demonstrated that *VEGF* levels are increased by several ADs and that VEGF signaling is required for AD-induced behavioral response [127].

In MDD patients, higher expression of *VEGF* mRNA in peripheral leukocytes was associated with a depressive state, and recovery of these levels occurred in parallel with clinical improvement [128]. Although it seems like a good candidate gene, *VEGF* has been poorly studied in AD pharmacogenetics, and markers rs1570360, rs2010963, rs25648, rs833069, rs3025010, rs3025033, and rs3025039 have provided negative results [129].

Dystrobrevin binding protein 1 (*DTNBP1* gene) may contribute to actin cytoskeleton reorganization, neurite outgrowth, synaptic signaling, regulation of glutamatergic neurotransmission [130], and neuroplasticity. This biological function may indicate that *DTNBP1* is promising as a candidate gene in AD response variability, but preliminary data are still difficult to interpret. Indeed, both negative and positive findings were reported for several haplotypes [12].

The complexity of signals produced by neurotrophic factors is mediated through signal transduction proteins, from the cell surface to the cytoplasm and nucleus. The guanine nucleotide binding protein (G protein) beta polypeptide 3 (*GNB3* gene) is involved in generation of second messengers in response to a number of signals—growth factors, neurotransmitters, hormones, and the like. Given the high complexity of G protein signal transduction cascades and their high diffusion, they are thought to be involved in neuronal plasticity [131] as well as in AD response. The most promising variant in *GNB3* is rs5443 (C825T), with a T allele that was associated with the occurrence of a splice variant (Gbeta3s) that appears to have altered activity [132]. Several independent studies found that the T allele predicted better AD response, but some data suggested an opposite effect especially in young subjects [12]. Negative findings [133–135] in Asian samples suggest a possible ethnic stratification.

Glycogen synthase kinase 3 alpha (*GSK3A*) acts as a regulator of Wnt/beta-catenin signaling, which is involved in the control of gene expression, cell behavior, cell adhesion, and cell polarity and plays a major role in neurodevelopment, regulation of neuroplasticity and cell survival [136]. The inhibitory phosphorylation of GSK3 occurs in the context of the signaling cascades in response to 5-HT, 5-HT1 receptor agonists, lithium, and ADs [137]. In particular, *GSK3A* may be an important mediator of 5-HT action in the brain and so of depressive-like behaviors [138]. The T allele of the promoter rs334558 (−50 T/C) was associated with transcriptional strength [139]; in some studies involving bipolar depressed patients different associations were shown depending on type of treatment [140,141]. In patients treated with only AD drugs, the C allele resulted in the favorable variant [142]. Given the interesting biological role of this gene from the perspective of AD mechanisms of action, further investigation will be a promising challenge.

Finally, the *CREB* gene (cAMP-responsive element binding protein) encodes a transcription factor that binds to the cAMP-responsive element and induces gene transcription in response to stimulation of the cAMP pathway. Alternate splicing of this gene results in two transcript variants encoding different isoforms, of which only *CREB1* was studied in relation to AD pharmacogenetics.

CREB activation is involved in proliferation of neuronal precursors, outgrowth of neuronal processes, learning and memory, induction of neurotrophic cellular programs (e.g., transcription of *BDNF*), and regulation of circadian rhythms. In humans, alterations in *CREB* have been associated with the pathophysiology of depression and suicide [143], as well as with the mechanism of AD action and response. Indeed, increased *CREB* levels in rodent models result in antidepressant-like behaviors, and studies on both humans and rodents have shown that *CREB* is upregulated by chronic AD treatment [144]. Thus, it can be hypothesized that specific variants in *CREB1* might be related to a lower likelihood of recovery from MDD, possibly through a differential modulation of gene expression and activation. Despite promising molecular and preclinical findings, *CREB1* has yet to be comprehensively investigated, but preliminary findings suggested its association with treatment-resistant MDD [145]. Future pharmacogenetic investigations of *CREB1* should take into account the number of failed antidepressant treatments.

26.2.2 Antidepressant Pharmacokinetics

Pharmacokinetics refers to all biological processes involved in drug absorption, distribution, transport, metabolism (e.g., by metabolic enzymes such as *cytochrome P450* or *glucuronosyltransferase* enzymes), and excretion from the body.

AD pharmacokinetics may be modified not only by enzyme inhibitors/enhancers, aging, and pregnancy, but also by polymorphisms located in key genes involved in AD metabolism and clearance. The genes coding for P-glycoprotein (P-gp) and cytochrome P450 (CYP) isoenzymes are the most interesting from this point of view because their products play a pivotal role in AD transport and metabolism [58].

P-glycoprotein (the *ABCB1* gene) is an ATP-dependent drug efflux pump for xenobiotic compounds that decreases drug accumulation in multidrug-resistant cells and limits uptake of some lipophilic drugs into key organs such as the brain. Animal studies show that a wide variety of structurally unrelated drugs are efficaciously carried out of the brain by P-gp activity, among which are a long list of ADs (with some exceptions, such as fluoxetine and bupropion) [58]. ABC transporters have been implicated in resistance to pharmacotherapy, most notably in oncology, where the expression of these efflux pumps by cancer cells can confer resistance to chemotherapy [146]. Over 50 functional SNPs have been identified in the *ABCB1* gene to date, and several studies demonstrated that they can impact P-gp expression and function in humans and therefore influence the pharmacokinetics of various drugs, including ADs [146]. The most studied polymorphisms in this gene are rs2032582 and rs1045642, since they alter P-gp expression and/or function [147], but their association with AD efficacy is still controversial [58].

Given that P-glycoprotein activity regulates drug uptake into tissues, *ABCB1* gene variations may also affect the risk of AD-induced side effects. However, this topic has not so far been extensively studied. The most systematic study investigated the correlation between *ABCB1* polymorphisms and the most common SSRI-induced side effects [148]. The main finding was the association between SSRI-related side effects and rs2032583 and rs2235040. Consistent preliminary evidence suggests that genetic variation in the *ABCB1* gene might influence the risk of switching and discontinuation of AD therapy [149]. Additional data indicate that the AD dose needed for remission may be predicted by a variant in the *ABCB1* gene [150].

The cytochrome P450 (CYP) superfamily is a class of enzymes that have a major effect on the oxidation and reduction of both endogenous and xenobiotic substances. The key role of these enzymes in drug metabolism and the highly polymorphic nature of their coding genes make them an interesting subject in pharmacogenetics. The most relevant CYP isoforms for AD metabolism are *CYP2D6*, *CYP2C19*, *CYP2C9*, *CYP1A2*, and *CYP2B6*. The known alleles show normal, reduced/absent, or increased activity (for a complete overview, see http://www.cypalleles.ki.se/), allowing some metabolizing groups to be distinguished.

In any case, this classification does not exactly reflect metabolic status because it does not take into account that some alleles are only partially active and some show different activity depending on the drug metabolized [151].

According to this classification, the wild-type genotype is defined as an extensive metabolizer (EM), characterized by the presence of two active alleles; an intermediate metabolizer (IM) is characterized by the presence of one wild-type allele plus a partially or totally defective allele and is expected to be between the EMs and the poor metabolizers (PMs), which show a combination of two partially or totally defective alleles. Finally, the ultrarapid metabolizer (UM) category exists only for *CYP2D6* and is usually linked to multiple allele copies. Metabolizing status was demonstrated to be associated with AD pharmacokinetics in a number of studies [12]—in particular, the *CYP2D6* gene and venlafaxine, fluoxetine, paroxetine, and nortriptyline pharmacokinetics; and the *CYP2C19* gene and citalopram, escitalopram, and amitriptyline/nortriptyline pharmacokinetics. On the basis of metabolizing status, theoretical dose adjustments were determined [152].

Available data do not support a correlation between AD plasma levels and the response of most ADs (with the exception of TCAs), and this is probably linked to the lack of association between response and CYP450 genetic polymorphisms found for the most part by previous studies [151]. Therefore, evidence does not as yet exist to support the recommendation of CYP genotyping in clinical practice, since the effect of CYP variants on clinical outcomes is still not completely clear for the most studied isoenzymes.

In conclusion, even if genotyping technologies and dose adjustments based on pharmacokinetic findings are already available, further investigation is expected to clarify their usefulness and applicability in clinical practice. Our current knowledge suggests that genotyping might be relevant for choosing a drug when the parent drug or its metabolites are critical for drug safety, and it may be that polymorphic drug metabolism affects the risk of side effects (as for TCAs).

26.3 GENOME-WIDE APPROACHES

In the following subsections, genome-wide association studies (GWAS) and some statistical methodologies based on genome-wide genotyping are discussed. GWAS showed rapid spread in genetic research during the last years, and the improvement of this technique represents a pivotal step in the dissection of antidepressant pharmacogenetics.

26.3.1 Genome-Wide Association Studies

Genome-wide association studies (GWAS) have been proposed as a powerful means of investigating human complex traits and have rapidly gained enthusiastic approval. GWAS provide some advantages compared to candidate gene studies. Indeed, they allow the genotyping of hundreds of thousands of polymorphisms across the whole genome and they overcome the need for a priori hypothesis. The hypothesis-free approach has the advantage of interrogating genes that have not been considered candidates because of our limited understanding of their function. Further, GWAS can uncover important regulatory variations in the large noncoding regions of the genome.

On the other hand, several limitations have to be considered, including the following:

- The available microarray platforms allow the genotyping of less than 50% of genomic variants [153].
- GWAS technology does not allow the reliable genotyping of rare variants (<1% of the population).
- There is the risk of the so called flip-flop phenomenon across different ethnic groups [154] because the functional gene effect is largely unknown and polymorphisms apparently associated with outcomes may only be in LD with truly associated variants.
- It is difficult to balance the risk of false-positive and false-negative findings. Indeed, the risk of false-positive findings should be carefully considered at least when data are unbalanced [155].

Regarding the final item in the list, the risk of a type I error can be very high when there are fewer than 100 observations in the minor genotype group, and this limitation cannot be overcome by increasing the total sample size. On the other hand, at the stringent genome-wide significance threshold, markers that are truly but weakly related to the phenotype may not be detected [156]. Indeed, the significance level is usually set between 1E-7 and 5E-8 [157] because of the very low prior probability of association and the limited number of variants determining the phenotype compared to the total number of markers genotyped. Prior information about genome regions that are related to the phenotype of interest may be used to boost the power of detecting signals, with a consequent balance between the risk of type I and type II errors [158].

Three large GWASs—the Genome-Based Therapeutic Drugs for Depression study (GENDEP) [119], the Sequenced Treatment Alternatives to Relieve Depression study (STAR*D), [159], and the Munich Antidepressant Response Signature study (MARS) [160]—have studied the effect of genetic variants on AD efficacy. In none of these were convincing genome-wide results reported.

In the STAR*D GWAS, among the top findings were the ubiquitin protein ligase E3C (*UBE3C*), the bone morphogenic protein 7 (*BMP7*), and the RAR-related orphan receptor A (*RORA*) genes. The *RORA* gene is related to circadian rhythm regulation [161], which is often altered in MDD. *UBE3C* is involved in protein degradation, while *BMP7*, a member of the transforming growth factor-beta (*TGFB*) superfamily, induces endochondral osteogenesis. For both these genes, a putative biological role in AD response is not an immediate interpretation. The same can be stated for Acyl-CoA synthetase short-chain family member 3 (*ACSS3*), a gene that may be associated with unsustained AD response [162]. Regarding *BMP7*, some kind of replication was obtained through full-exome sequencing, a method based on a similar GWAS principle but focused on exons (i.e., coding regions) only. The finding involved the bone morphogenetic protein 5 (*BMP5*) [163], which is a member of the same family as *BMP7*.

The MARS study retrieved the top signals in the cadherin-17 (*CDH17*) and the EPH receptor B1 (*EPHB1*) genes. Cadherins are calcium-dependent cell adhesion proteins that may play an important role in synapse formation in the hippocampus [164]—a brain region involved in MDD pathogenesis. *EPHB1* is an ephrin receptor, and in adults regulates chemotaxis, proliferation, and polarity of the hippocampus neural progenitors.

The GENDEP study produced interesting results for the interleukin-11 (*IL11*) and uronyl 2-sulphotransferase (*UST*) genes. The finding in the *IL11* gene appears to confirm the inflammatory hypothesis of MDD and the anti-inflammatory properties of ADs. *UST* codes for an enzyme that transfers sulfate to uronyl residues of glycosaminoglycans (components of the extracellular matrix). A collaborative analysis by STAR*D and GENDEP showed the strongest common signal from the chondroitin sulfate N-acetylgalactosaminyltransferase 1 gene (*CSGALNACT1*) [162], which is involved in chondroitin chain biosynthesis in cartilage. This finding may be interesting in view of the signal found in the *UST* gene. Although chondroitin sulfate has a well established role in the skeletal system, it may also play a part in neuroplasticity, regeneration, and brain development [165].

AD efficacy is influenced by environmental factors that interact with genes [166]. Genome-wide data, as well as candidate gene data, have been investigated with the aim of retrieving meaningful gene × environment interactions. In the STAR*D study, the neural precursor cell expressed developmentally downregulated 4-like, E3 ubiquitin protein ligase (*NEDD4L*) gene interacts with patient quality of life in determining AD response [167]. Interestingly, this finding confirmed the hypothesis of involvement of ubiquitin ligases in AD response (see STAR*D findings reported on the previous page).

The most significant GWAS findings diverge across different samples, mainly because of the heavy influence of stratification factors (such as cryptic population stratification) on the typical small effect sizes detected by GWAS. To increase statistical power, recently GWAS meta-analyses have been performed, albeit that they increase signal divergence because of intersample stratification. Accordingly, no significant findings have been reported using this approach [168]. On the other hand, analyses focused on more homogeneous groups in terms of treatment revealed a concordant signal among the main GWAS previously cited, but they explain only ≈1.2% of the variance in response [169].

In addition to therapeutic effects, one study investigated genome-wide data to identify markers of AD-induced side effects. In detail, the SAC1 suppressor of actin mutations 1-like (yeast) (*SACM1L*) gene revealed itself as a putative modulator of bupropion-related sexual side effects [170]. *SACM1L* encodes an integral membrane protein of the endoplasmic reticulum and the Golgi apparatus that influences vesicle trafficking; *SACM1L* shows wide tissue distribution and mutations produce a wide array of phenotypes. This study found that the gene *EMID2* (EMI domain containing 2) affects vision/hearing side effects [171]. *EMID2* has a possible biological rationale for association with these side effects, since it encodes the protein collagen α-1 chain involved in the regulation of corneal collagen fibrillogenesis.

Despite these findings, no clinical applications are expected in the next years given the small effect size of the results, the lack of replication, and the absence of cost−benefit studies. For a summary of GWAS results see Table 26.2.

26.3.2 Pathway Analysis

Recently pathway analysis was proposed as a way to balance some of the limitations of hypothesis-free GWAS. Pathway analysis looks at gene variants that are involved in the same biological pathway. Thus, genome-wide data can be analyzed as units of biologically related genes and not only as single genetic variants. Pathway analysis may yield more insight into disease biology because it overcomes the genetic heterogeneity bias (e.g., due to population stratification and differential rates of genotyping error between the groups under analysis). Also, setting the variants in the same pathway as the unit of analysis may increase the power needed to detect associations and replicate findings [172].

Indeed, although individual SNPs did not reach significance in many GWASs that investigated the genetics of MDD, ranking SNPs previously associated with different psychiatric disorders identified pathway convergence in synaptogenesis, axonal guidance, synaptic plasticity, and calcium signaling [173]. Similarly, some main molecular clusters were associated with AD efficacy. In detail, genes that belong to metabolic pathways and brain development, metabolic and cardiovascular disease, and GABAergic and glutamatergic neurotransmission are putative modulators of AD response [160]. Among genes involved in brain development, the strongest signals were found in the transcription factor nuclear receptor subfamily 2, group E, member 1 (*NR2E1*), the substrate (ephrin-A5; *EFNA5*), and the receptor (EPH receptor A5; *EPHA5*) genes of ephrin A5, which is an important modulator of late-stage nervous system development and differentiation.

The metabolic and cardiovascular disease gene cluster was so named because it includes potential risk genes for cardiovascular disorders, such as CD36 paraoxonase 2 (*PON2*); apolipoprotein B (*APOB*); and phosphoinositide-3-kinase, regulatory subunit 1 (*PIK3R1*). Abnormalities in glucose and lipid metabolism, as well as cellular antioxidant potential, can clearly also influence neuronal survival and activity, particularly synaptic plasticity, and finally impact on AD response.

The gene group related to GABAergic and glutamatergic neurotransmission includes, among others, neuregulin 1(*NRG1*) and homer homolog 1 (*HOMER1*), which are involved in neurodevelopment and neuroplasticity; and glial high-affinity glutamate transporter (*SLC1A2*), which clears glutamate from synapses and can prevent cellular damage from excessive activation of glutamate receptors. The central glutamatergic system represents a key regulator of long-term potentiation, a process that underlies multiple forms of long-term memory in the mammalian brain. Recently, the involvement of the long-term potentiation pathway in AD response was confirmed [162].

The described gene clusters show a clear biological rationale for involvement in AD efficacy, confirming the substantial discovery potential of pathway analysis compared to methods based on the analysis of individual markers.

26.4 CONCLUSION

As reported in this chapter, a high number of pharmacokinetic and pharmacodynamic genes may contribute to the clinical outcomes of AD treatment and explain the hypothesized 50% variance in response ascribed to genetic factors. Nonetheless, pharmacogenetic studies have so far provided poorly replicated and often contradictory results. A possible explanation is the main phenotype investigated so far: MDD itself. Indeed, standard criteria for the diagnosis of MDD are probably not discriminative enough to identify homogeneous patient groups.

There is increasing evidence that MDD represents a common final phenotype deriving from various pathophysiological mechanisms. Therefore, in recent years more homogeneous patient subgroups characterized by common features (i.e., endophenotypes) have been studied. The endophenotype is a concept characterized by specific properties. In detail: (1) the endophenotype is heritable; (2) the endophenotype is primarily state independent (i.e., it manifests in an individual whether or not illness is active); (3) in families, endophenotype and illness cosegregate; and (4) the endophenotype found in affected family members is found in nonaffected family members at a higher rate than in the general population. Thus, endophenotypes are measureable characteristics that fill the gaps between available descriptors and those between the gene and the elusive disease process. Behavioral and neural factors and biomarkers have been tested for their significance in defining MDD endophenotypes [174].

Current and future pharmacogenetic studies are expected to collect detailed data regarding not only major clinical features but also neuroimaging and neuropsychological phenotypes, thus allowing the identification of depressive endophenotypes. Other factors involved in poor replication of pharmacogenetic findings are probably ethnic stratification, poor covering of genetic variance by candidate gene studies, and interaction between genetic and environmental factors. Other issues are represented by the so-called flip-flop phenomenon (investigation of a noncausal variant in LD with a genuine variant [154]), de novo mutations, and epigenetic modifications. Indeed, preliminary findings suggest that epigenetic mechanisms may play a role in regulating the expression of key genes during AD treatment [175]. On the other hand, the effect of de novo (point substitutions or single-nucleotide variants and small insertions or deletions) and recurrent mutations has been hypothesized as a further cause of poor replication of genetic findings.

Regarding pharmacokinetic genes, other specific limitations have to be considered. First, a single drug is metabolized by specific enzymes, but the result of this reaction may be an active molecule that can be metabolized by one or more different enzymes. Second, drug interactions may strongly affect the cytochrome activity in a way that may be dependent on gene variants but is, to an extent, still difficult to define. For example, the enzyme activity of a genetic PM cannot be affected by inducers or inhibitors, but EMs or UMs can be converted to the PM phenotype by strong enzyme inhibitors. Finally, drugs can be metabolized by a set of cytochromes: they should all be investigated by a single test in order to infer sufficient predictive information about drug blood levels.

Assuming that these limitations can be overcome by better study design and technological improvements, there should be a multistep process for identifying genetic marker sets that may be useful in clinical practice (Figure 26.1). This process should include molecular studies, cell line experiments, animal models, neuroimaging, and biomarkers in MDD patients, all of which are expected to identify the biological mechanisms involved in AD effect. In this effort, GWASs can contribute to the discovery of molecules with still unknown function and intergenic markers. Candidate studies can confirm the involvement of marker sets in AD outcome thanks to whole-genome sequencing. Selected genes can thus be grouped into molecular pathways that can be tested as a single unit for their effect on AD response. Finally, cost-effectiveness trials on the identified gene sets should verify the actual applicability of genotyping in clinical practice. This is clearly only one method for identifying genetic markers; several complementary and partially overlapping methods can be used to reach the same goal.

Studies in the multistep discovery process of genetic markers have been performed, even if some steps have yet to be sufficiently analyzed (in particular, pathway and

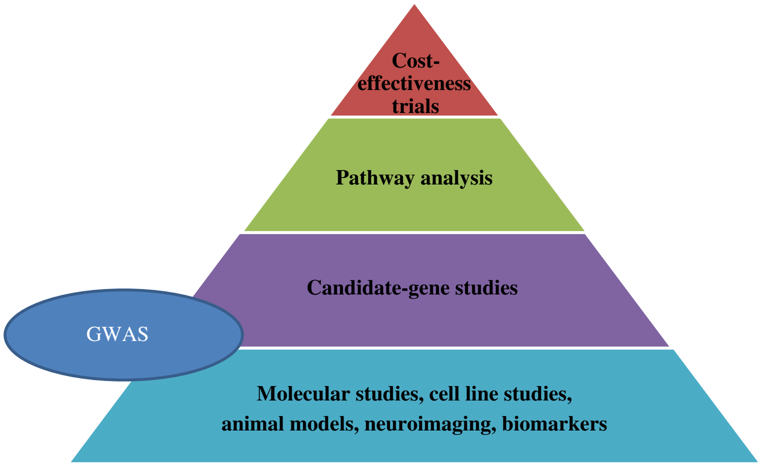

Cost-effectiveness trials
Pathway analysis
Candidate-gene studies
GWAS
Molecular studies, cell line studies, animal models, neuroimaging, biomarkers

FIGURE 26.1 **Multistep identification of sets of genetic markers useful in clinical practice.** Shown at the bottom of the pyramid are studies aimed at identifying the biological mechanisms involved in AD action. GWASs attempt to discover molecules with as yet unknown function. Candidate studies can confirm the involvement of gene sets in AD outcomes thanks to whole-genome sequencing. Selected genes can thus be grouped into molecular pathways that can be tested as a single unit for their effect on antidepressant response. Finally, cost-effectiveness trials on the identified gene sets should verify the actual applicability of genotyping in clinical practice.

cost-effectiveness analysis). Nevertheless, some data about the cost-effectiveness of genotyping already exist. Indeed, the genotyping of 5-HTTLPR prior to AD treatment was associated with an increase both in AD response and tolerability compared to nongenotyping. Further, genotyping shows a positive cost-effectiveness profile after at least two recurrent MDD episodes [176]. In addition, a study of cost-effectiveness data from 27 European states regarding AD choice based on response and tolerability predicted by 5-HTTLPR profile compared to an usual algorithm has been performed [177]. It found the probability of an incremental cost-effectiveness ratio value below the WHO-recommended cost-utility threshold to be >90% in high-income countries.

Finally, we underline that pharmacogenetics provides a unique tool for increasing knowledge of AD mechanisms of action. The idea that unsatisfactory AD response rates are partially due to the lack of drugs with alternative mechanism of action is well supported [178]. Thus, the engineering of molecules with truly innovative mechanisms of action appears to be a primary objective together with the identification of response predictors.

KEY CONCEPTS

- The estimated variance in antidepressant response due to genetic variants is around 50%
- Results of candidate gene studies were often contradictory and poorly replicated, but outlined few good candidate genes (SLC6A4, COMT, IL1B, BDNF, and GNB3)
- Genome-wide association studies provide pivotal data to improve the limited knowledge derived by candidate gene studies, and suggested RORA, CDH17, EPHB1, and IL11 as promising new candidate genes
- A multistep and convergent pharmacogenetic approach could provide more informative data than genome-wide association studies and candidate gene studies used as independent tools
- Preliminary cost-effectiveness trials supported the applicability of genotyping prior to antidepressant treatment beginning, but neither approved genetic test nor specific clinical guidelines to use it still exist

GLOSSARY

Genome-wide association study the examination of hundreds of thousands of common genetic variants to study associations with human complex traits.

Linkage disequilibrium the nonrandom association of alleles at two or more loci that descend from single, ancestral chromosomes.

Phenotype an organism's observable characteristic or trait, such as biochemical or physiological properties or behavior. Phenotypes result from the expression of an organism's genes as well as the influence of environmental factors and the interactions between the two.

Polymorphism the natural variation in a DNA sequence that has no evident adverse effects and occurs with fairly high frequency in the general population.

SNP the DNA sequence variation that occurs when a single nucleotide (A, T, C, or G) in the genome differs among members of a biological species or among paired human chromosomes.

VNTR the location in a genome where a short nucleotide sequence is organized as a tandem repeat, often showing variations in length between individuals.

ACRONYMS AND ABBREVIATIONS

AD	antidepressant
GWAS	genome-wide association study
LD	linkage disequilibrium
MDD	major depressive disorder
SNP	single-nucleotide polymorphism
VNTR	variable number of tandem repeats

REFERENCES

[1] Warden D, Rush AJ, Trivedi MH, Fava M, Wisniewski SR. The STAR*D project results: a comprehensive review of findings. Curr Psychiatry Rep 2007;9(6):449–59.

[2] Trivedi MH, Daly EJ. Treatment strategies to improve and sustain remission in major depressive disorder. Dialogues Clin Neurosci 2008;10(4):377–84.

[3] Porcelli S, Drago A, Fabbri C, Gibiino S, Calati R, Serretti A. Pharmacogenetics of antidepressant response. J Psychiatry Neurosci 2011;36(2):87–113.

[4] Franchini L, Serretti A, Gasperini M, Smeraldi E. Familial concordance of fluvoxamine response as a tool for differentiating mood disorder pedigrees. J Psychiatr Res 1998;32(5):255–9.

[5] Porcelli S, Drago A, Fabbri C, Serretti A. Mechanisms of antidepressant action: an integrated dopaminergic perspective. Prog Neuropsychopharmacol Biol Psychiatry 2011;35(7):1532–43.

[6] Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry 2000;157(10):1552–62.

[7] Vesell ES. Pharmacogenetic perspectives gained from twin and family studies. Pharmacol Ther 1989;41(3):535–52.

[8] Roberts R, Wells GA, Stewart AF, Dandona S, Chen L. The genome-wide association study–a new era for common polygenic disorders. J Cardiovasc Transl Res 2010;3(3):173–82.

[9] Nutt D, Demyttenaere K, Janka Z, Aarre T, Bourin M, Canonico PL, et al. The other face of depression, reduced positive affect: the role of catecholamines in causation and cure. J Psychopharmacol 2007;21(5):461–71.

[10] Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D, et al. Allelic variation of human serotonin transporter gene expression. J Neurochem 1996;66(6):2621–4.

[11] Scharinger C, Rabl U, Pezawas L, Kasper S. The genetic blueprint of major depressive disorder: contributions of imaging genetics studies. World J Biol Psychiatry 2011;12(7):474–88.

[12] Fabbri C, Di Girolamo G, Serretti A. Pharmacogenetics of antidepressant drugs: an update after almost 20 years of research. Am J: the official publication of the international society of psychiatric Med Genet B Neuropsychiatr Genet 2013;162B(6):487–520.

[13] Kunugi H, Hattori M, Kato T, Tatsumi M, Sakai T, Sasaki T, et al. Serotonin transporter gene polymorphisms: ethnic difference and possible association with bipolar affective disorder. Mol Psychiatry 1997;2(6):457–62.

[14] Porcelli S, Fabbri C, Serretti A. Meta-analysis of serotonin transporter gene promoter polymorphism (5-HTTLPR) association with antidepressant efficacy. Eur Neuropsychopharmacol 2012;22(4):239–58.

[15] Hu XZ, Lipsky RH, Zhu G, Akhtar LA, Taubman J, Greenberg BD, et al. Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. Am J Hum Genet 2006;78(5):815–26.

[16] Ruhe HG, Ooteman W, Booij J, Michel MC, Moeton M, Baas F, et al. Serotonin transporter gene promoter polymorphisms modify the association between paroxetine serotonin transporter occupancy and clinical response in major depressive disorder. Pharmacogenet Genomics 2009;19(1):67–76.

[17] Perroud N, Salzmann A, Saiz PA, Baca-Garcia E, Sarchiapone M, Garcia-Portilla MP, et al. Rare genotype combination of the serotonin transporter gene associated with treatment response in severe personality disorder. Am J Med Genet B Neuropsychiatr Genet 2010;153B(8):1494–7.

[18] Hranilovic D, Stefulj J, Schwab S, Borrmann-Hassenbach M, Albus M, Jernej B, et al. Serotonin transporter promoter and intron 2 polymorphisms: relationship between allelic variants and gene expression. Biol Psychiatry 2004;55(11):1090–4.

[19] Fiskerstrand CE, Lovejoy EA, Quinn JP. An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enhancer activity in embryonic stem cells. FEBS Lett 1999;458(2):171–4.

[20] MacKenzie A, Quinn J. A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. Proc Natl Acad Sci USA 1999;96(26):15251–5.

[21] Kato M, Serretti A. Review and meta-analysis of antidepressant pharmacogenetic findings in major depressive disorder. Mol Psychiatry 2010;15(5):473–500.

[22] Perez V, Gilaberte I, Faries D, Alvarez E, Artigas F. Randomised, double-blind, placebo-controlled trial of pindolol in combination with fluoxetine antidepressant treatment. Lancet 1997;349(9065):1594–7.

[23] Zhao X, Huang Y, Li J, Ma H, Jin Q, Wang Y, et al. Association between the 5-HT1A receptor gene polymorphism (rs6295) and antidepressants: a meta-analysis. Int Clin Psychopharmacol 2012;27(6):314–20.

[24] Peters EJ, Slager SL, McGrath PJ, Knowles JA, Hamilton SP. Investigation of serotonin-related genes in antidepressant response. Mol Psychiatry 2004;9(9):879–89.

[25] Illi A, Setala-Soikkeli E, Viikki M, Poutanen O, Huhtala H, Mononen N, et al. 5-HTR1A, 5-HTR2A, 5-HTR6, TPH1 and TPH2 polymorphisms and major depression. Neuroreport 2009;20(12):1125–8.

[26] Noro M, Antonijevic I, Forray C, Kasper S, Kocabas NA, Lecrubier Y, et al. 5HT1A and 5HT2A receptor genes in treatment response phenotypes in major depressive disorder. Int Clin Psychopharmacol 2010;25(4):228–31.

[27] Serretti A, Fabbri C, Pellegrini S, Porcelli S, Politi P, Bellino S, et al. No effect of serotoninergic gene variants on response to interpersonal counseling and antidepressants in major depression. Psychiatry Invest 2013;10(2):180–9.

[28] Yu YW, Tsai SJ, Liou YJ, Hong CJ, Chen TJ. Association study of two serotonin 1A receptor gene polymorphisms and fluoxetine treatment response in Chinese major depressive disorders. Eur Neuropsychopharmacol 2006;16(7):498–503.

[29] Baune BT, Hohoff C, Roehrs T, Deckert J, Arolt V, Domschke K. Serotonin receptor 1A–1019C/G variant: impact on antidepressant pharmacoresponse in melancholic depression? Neurosci Lett 2008;436(2):111–5.

[30] Suzuki Y, Sawamura K, Someya T. The effects of a 5-hydroxytryptamine 1A receptor gene polymorphism on the clinical response to fluvoxamine in depressed patients. Pharmacogenomics J 2004;4(4):283–6.

[31] Levin GM, Bowles TM, Ehret MJ, Langaee T, Tan JY, Johnson JA, et al. Assessment of human serotonin 1A receptor polymorphisms and SSRI responsiveness. Molecular Diagn Ther 2007;11(3):155–60.

[32] Dhaenen H. Imaging the serotonergic system in depression. Eur Arch Psychiatry Clin Neurosci 2001;251(Suppl 2):II76–80.

[33] McMahon FJ, Buervenich S, Charney D, Lipsky R, Rush AJ, Wilson AF, et al. Variation in the gene encoding the serotonin 2A receptor is associated with outcome of antidepressant treatment. Am J Hum Genet 2006;78(5):804–14.

[34] Peters EJ, Slager SL, Jenkins GD, Reinalda MS, Garriock HA, Shyn SI, et al. Resequencing of serotonin-related genes and association of tagging SNPs to citalopram response. Pharmacogenet Genomics 2009;19(1):1–10.

[35] Perlis RH, Fijal B, Adams DH, Sutton VK, Trivedi MH, Houston JP. Variation in catechol-O-methyltransferase is associated with duloxetine response in a clinical trial for major depressive disorder. Biol Psychiatry 2009;65(9):785–91.

[36] Uher R, Huezo-Diaz P, Perroud N, Smith R, Rietschel M, Mors O, et al. Genetic predictors of response to antidepressants in the GENDEP project. Pharmacogenomics J 2009;9(4):225–33.

[37] Kishi T, Yoshimura R, Kitajima T, Okochi T, Okumura T, Tsunoka T, et al. HTR2A is associated with SSRI response in major depressive disorder in a Japanese cohort. Neuromolecular Med 2010;12(3):237–42.

[38] Horstmann S, Lucae S, Menke A, Hennings JM, Ising M, Roeske D, et al. Polymorphisms in GRIK4, HTR2A, and FKBP5 show interactive effects in predicting remission to antidepressant treatment. Neuropsychopharmacology 2010;35(3):727–40.

[39] Lucae S, Ising M, Horstmann S, Baune BT, Arolt V, Muller-Myhsok B, et al. HTR2A gene variation is involved in antidepressant treatment response. Eur Neuropsychopharmacol 2010;20(1):65–8.

[40] Viikki M, Huuhka K, Leinonen E, Illi A, Setala-Soikkeli E, Huuhka M, et al. Interaction between two HTR2A polymorphisms and gender is associated with treatment response in MDD. Neurosci Lett 2011;501(1):20–4.

[41] Lin E, Chen PS, Chang HH, Gean PW, Tsai HC, Yang YK, et al. Interaction of serotonin-related genes affects short-term antidepressant response in major depressive disorder. Prog Neuropsychopharmacol Biol Psychiatry 2009;33(7):1167–72.

[42] Spurlock G, Heils A, Holmans P, Williams J, D'Souza UM, Cardno A, et al. A family based association study of T102C polymorphism in 5HT2A and schizophrenia plus identification of new polymorphisms in the promoter. Mol Psychiatry 1998;3(1):42–9.

[43] Paddock S, Laje G, Charney D, Rush AJ, Wilson AF, Sorant AJ, et al. Association of GRIK4 with outcome of antidepressant treatment in the STAR*D cohort. Am J Psychiatry 2007;164(8):1181–8.

[44] Yoshida K, Naito S, Takahashi H, Sato K, Ito K, Kamata M, et al. Monoamine oxidase: a gene polymorphism, tryptophan hydroxylase gene polymorphism and antidepressant response to fluvoxamine in Japanese patients with major depressive disorder. Prog Neuropsychopharmacol Biol Psychiatry 2002;26(7–8):1279–83.

[45] Murphy Jr GM, Kremer C, Rodrigues HE, Schatzberg AF. Pharmacogenetics of antidepressant medication intolerance. Am J Psychiatry 2003;160(10):1830–5.

[46] Tanaka M, Kobayashi D, Murakami Y, Ozaki N, Suzuki T, Iwata N, et al. Genetic polymorphisms in the 5-hydroxytryptamine type 3B receptor gene and paroxetine-induced nausea. Int J Neuropsychopharmacol 2008;11(2):261–7.

[47] Higuchi H, Takahashi H, Kamata M, Yoshida K. Influence of serotonergic/noradrenergic gene polymorphisms on nausea and sweating induced by milnacipran in the treatment of depression. Neuropsychiatr Dis Treat 2009;5:393–8.

[48] Kato M, Fukuda T, Wakeno M, Fukuda K, Okugawa G, Ikenaga Y, et al. Effects of the serotonin type 2A, 3A and 3B receptor and the serotonin transporter genes on paroxetine and fluvoxamine efficacy and adverse drug reactions in depressed Japanese patients. Neuropsychobiology 2006;53(4):186–95.

[49] Sugai T, Suzuki Y, Sawamura K, Fukui N, Inoue Y, Someya T. The effect of 5-hydroxytryptamine 3A and 3B receptor genes on nausea induced by paroxetine. Pharmacogenomics J 2006;6(5):351–6.

[50] Ballaz SJ, Akil H, Watson SJ. Analysis of 5-HT6 and 5-HT7 receptor gene expression in rats showing differences in novelty-seeking behavior. Neuroscience 2007;147(2):428–38.

[51] Mitchell ES, Sexton T, Neumaier JF. Increased expression of 5-HT6 receptors in the rat dorsomedial striatum impairs instrumental learning. Neuropsychopharmacology 2007;32(7):1520–30.

[52] Svenningsson P, Tzavara ET, Qi H, Carruthers R, Witkin JM, Nomikos GG, et al. Biochemical and behavioral evidence for antidepressant-like effects of 5-HT6 receptor stimulation. J Neurosci 2007;27(15):4201–9.

[53] Kohen R, Metcalf MA, Khan N, Druck T, Huebner K, Lachowicz JE, et al. Cloning, characterization, and chromosomal localization of a human 5-HT6 serotonin receptor. J Neurochem 1996;66(1):47–56.

[54] Lee SH, Lee KJ, Lee HJ, Ham BJ, Ryu SH, Lee MS. Association between the 5-HT6 receptor C267T polymorphism and response to antidepressant treatment in major depressive disorder. Psychiatry Clin Neurosci 2005;59(2):140–5.

[55] Jonsson EG, Goldman D, Spurlock G, Gustavsson JP, Nielsen DA, Linnoila M, et al. Tryptophan hydroxylase and catechol-O-methyltransferase gene polymorphisms: relationships to monoamine metabolite concentrations in CSF of healthy volunteers. Eur Arch Psychiatry Clin Neurosci 1997;247(6):297–302.

[56] Secher A, Bukh J, Bock C, Koefoed P, Rasmussen HB, Werge T, et al. Antidepressive-drug-induced bodyweight gain is associated with polymorphisms in genes coding for COMT and TPH1. Int Clin Psychopharmacol 2009;24(4):199–203.

[57] Anttila S, Huuhka K, Huuhka M, Rontu R, Mattila KM, Leinonen E, et al. Interaction between TPH1 and GNB3 genotypes and electroconvulsive therapy in major depression. J Neural Transm 2007;114(4):461–8.

[58] Porcelli S, Fabbri C, Drago A, Gibiino S, De Ronchi D, Serretti A. Genetics and antidepressants: where we are. Clin Neuropsychiatry 2011;7(2):99–150.

[59] Kocabas NA. Catechol-O-methyltransferase (COMT) pharmacogenetics in the treatment response phenotypes of major depressive disorder (MDD). CNS Neurol Disord Drug Targets 2012;11(3):264–72.

[60] Kocabas NA, Faghel C, Barreto M, Kasper S, Linotte S, Mendlewicz J, et al. The impact of catechol-O-methyltransferase SNPs and haplotypes on treatment response phenotypes in major depressive disorder: a case-control association study. Int Clin Psychopharmacol 2010;25(4):218–27.

[61] Ji Y, Biernacka J, Snyder K, Drews M, Pelleymounter LL, Colby C, et al. Catechol O-methyltransferase pharmacogenomics and selective serotonin reuptake inhibitor response. Pharmacogenomics J 2012;12(1):78–85.

[62] Dong C, Wong ML, Licinio J. Sequence variations of ABCB1, SLC6A2, SLC6A3, SLC6A4, CREB1, CRHR1 and NTRK2: association with major depression and antidepressant response in Mexican-Americans. Mol Psychiatry 2009;14(12):1105–18.

[63] Sabol SZ, Hu S, Hamer D. A functional polymorphism in the monoamine oxidase A gene promoter. Hum Genet 1998;103(3):273–9.

[64] Jonsson EG, Norton N, Gustavsson JP, Oreland L, Owen MJ, Sedvall GC. A promoter polymorphism in the monoamine oxidase A

gene and its relationships to monoamine metabolite concentrations in CSF of healthy volunteers. J Psychiatr Res 2000;34(3):239–44.

[65] Fowler JS, Alia-Klein N, Kriplani A, Logan J, Williams B, Zhu W, et al. Evidence that brain MAO A activity does not correspond to MAO A genotype in healthy male subjects. Biol Psychiatry 2007;62(4):355–8.

[66] Tzeng DS, Chien CC, Lung FW, Yang CY. MAOA gene polymorphisms and response to mirtazapine in major depression. Hum Psychopharmacol 2009;24(4):293–300.

[67] Yu YW, Tsai SJ, Hong CJ, Chen TJ, Chen MC, Yang CW. Association study of a monoamine oxidase a gene promoter polymorphism with major depressive disorder and antidepressant response. Neuropsychopharmacology 2005;30(9):1719–23.

[68] Domschke K, Hohoff C, Mortensen LS, Roehrs T, Deckert J, Arolt V, et al. Monoamine oxidase A variant influences antidepressant treatment response in female patients with major depression. Prog Neuropsychopharmacol Biol Psychiatry 2008;32(1):224–8.

[69] Keers R, Aitchison KJ. Gender differences in antidepressant drug response. Int Rev Psychiatry 2010;22(5):485–500.

[70] Crissman AM, Makhay MM, O'Donnell JM. Discriminative stimulus effects of centrally administered isoproterenol in rats: mediation by beta-1 adrenergic receptors. Psychopharmacology 2001;154(1):70–5.

[71] Zill P, Baghai TC, Engel R, Zwanzger P, Schule C, Minov C, et al. Beta-1-adrenergic receptor gene in major depression: influence on antidepressant treatment response. Am J Med Genet B Neuropsychiatr Genet 2003;120B(1):85–9.

[72] Lee HY, Kang RH, Paik JW, Jeong YJ, Chang HS, Han SW, et al. Association of the adrenergic alpha 2a receptor–1291C/G polymorphism with weight change and treatment response to mirtazapine in patients with major depressive disorder. Brain Res 2009;1262:1–6.

[73] Yanpallewar SU, Fernandes K, Marathe SV, Vadodaria KC, Jhaveri D, Rommelfanger K, et al. Alpha2-adrenoceptor blockade accelerates the neurogenic, neurotrophic, and behavioral effects of chronic antidepressant treatment. J Neurosci 2010;30(3):1096–109.

[74] Gillman PK. Tricyclic antidepressant pharmacology and therapeutic drug interactions updated. Br J Pharmacol 2007;151(6):737–48.

[75] Yoshida K, Takahashi H, Higuchi H, Kamata M, Ito K, Sato K, et al. Prediction of antidepressant response to milnacipran by norepinephrine transporter gene polymorphisms. Am J Psychiatry 2004;161(9):1575–80.

[76] Kim H, Lim SW, Kim S, Kim JW, Chang YH, Carroll BJ, et al. Monoamine transporter gene polymorphisms and antidepressant response in Koreans with late-life depression. JAMA 2006;296(13):1609–18.

[77] McClure DJ. The role of dopamine in depression. Can J Psychiatry 1973;18(4):309012.

[78] Willner P. Dopamine and depression: a review of recent evidence. I. Empirical studies. Brain Res 1983;287(3):211–24.

[79] D'Aquila PS, Collu M, Gessa GL, Serra G. The role of dopamine in the mechanism of action of antidepressant drugs. Eur J Pharmacol 2000;405(1–3):365–73.

[80] Willner P, Hale AS, Argyropoulos S. Dopaminergic mechanism of antidepressant action in depressed patients. J Affect Disord 2005;86(1):37–45.

[81] Serretti A, Zanardi R, Cusin C, Rossini D, Lilli R, Lorenzi C, et al. No association between dopamine D(2) and D(4) receptor gene variants and antidepressant activity of two selective serotonin reuptake inhibitors. Psychiatry Res 2001;104(3):195–203.

[82] Xu Z, Zhang Z, Shi Y, Pu M, Yuan Y, Zhang X, et al. Influence and interaction of genetic polymorphisms in catecholamine neurotransmitter systems and early life stress on antidepressant drug response. J Affect Disord 2011;133(1–2):165–73.

[83] Wang Y, Liu X, Yu Y, Han Y, Wei J, Collier D, et al. The role of single nucleotide polymorphism of D2 dopamine receptor gene on major depressive disorder and response to antidepressant treatment. Psychiatry Res 2012;200(2–3):1047–50.

[84] Breuer ME, Groenink L, Oosting RS, Buerger E, Korte M, Ferger B, et al. Antidepressant effects of pramipexole, a dopamine D3/D2 receptor agonist, and 7-OH-DPAT, a dopamine D3 receptor agonist, in olfactory bulbectomized rats. Eur J Pharmacol 2009;616(1–3):134–40.

[85] Garriock HA, Delgado P, Kling MA, Carpenter LL, Burke M, Burke WJ, et al. Number of risk genotypes is a risk factor for major depressive disorder: a case control study. Behav Brain Funct 2006;2:24.

[86] Zhou Z, Roy A, Lipsky R, Kuchipudi K, Zhu G, Taubman J, et al. Haplotype-based linkage of tryptophan hydroxylase 2 to suicide attempt, major depression, and cerebrospinal fluid 5-hydroxyindoleacetic acid in 4 populations. Arch Gen Psychiatry 2005;62(10):1109–18.

[87] Fuke S, Suo S, Takahashi N, Koike H, Sasagawa N, Ishiura S. The VNTR polymorphism of the human dopamine transporter (DAT1) gene affects gene expression. Pharmacogenomics J 2001;1(2):152–6.

[88] Kirchheiner J, Nickchen K, Sasse J, Bauer M, Roots I, Brockmoller J. A 40-basepair VNTR polymorphism in the dopamine transporter (DAT1) gene and the rapid response to antidepressant treatment. Pharmacogenomics J 2007;7(1):48–55.

[89] Sapolsky RM. The possibility of neurotoxicity in the hippocampus in major depression: a primer on neuron death. Biol Psychiatry 2000;48(8):755–65.

[90] Sanacora G, Mason GF, Rothman DL, Behar KL, Hyder F, Petroff OA, et al. Reduced cortical gamma-aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy. Arch Gen Psychiatry 1999;56(11):1043–7.

[91] Skolnick P, Popik P, Trullas R. Glutamate-based antidepressants: 20 years on. Trends Pharmacol Sci 2009;30(11):563–9.

[92] McEwen BS, Conrad CD, Kuroda Y, Frankfurt M, Magarinos AM, McKittrick C. Prevention of stress-induced morphological and cognitive consequences. Eur Neuropsychopharmacol 1997;7(Suppl 3):S323–8.

[93] Zhang H, Etherington LA, Hafner AS, Belelli D, Coussen F, Delagrange P, et al. Regulation of AMPA receptor surface trafficking and synaptic plasticity by a cognitive enhancer and antidepressant molecule. Mol Psychiatry 2013;18(4):471–84.

[94] Drago A, Crisafulli C, Sidoti A, Serretti A. The molecular interaction between the glutamatergic, noradrenergic, dopaminergic and serotoninergic systems informs a detailed genetic perspective on depressive phenotypes. Prog Neurobiol 2011;94(4):418–60.

[95] Catches JS, Xu J, Contractor A. Genetic ablation of the GluK4 kainate receptor subunit causes anxiolytic and antidepressant-like behavior in mice. Behav Brain Res 2012;228(2):406–14.

[96] Fabbri C, Drago A, Serretti A. Early antidepressant efficacy modulation by glutamatergic gene variants in the STAR*D. Eur Neuropsychopharmacol 2013;23(7):612–21.

[97] Perlis RH, Moorjani P, Fagerness J, Purcell S, Trivedi MH, Fava M, et al. Pharmacogenetic analysis of genes implicated in rodent

models of antidepressant response: association of TREK1 and treatment resistance in the STAR(*)D study. Neuropsychopharmacology 2008;33(12):2810–9.

[98] Reichenberg A, Yirmiya R, Schuld A, Kraus T, Haack M, Morag A, et al. Cytokine-associated emotional and cognitive disturbances in humans. Arch Gen Psychiatry 2001;58(5):445–52.

[99] Reichenberg A, Kraus T, Haack M, Schuld A, Pollmacher T, Yirmiya R. Endotoxin-induced changes in food consumption in healthy volunteers are associated with TNF-alpha and IL-6 secretion. Psychoneuroendocrinology 2002;27(8):945–56.

[100] Holsboer F. The corticosteroid receptor hypothesis of depression. Neuropsychopharmacology 2000;23(5):477–501.

[101] Janssen DG, Caniato RN, Verster JC, Baune BT. A psychoneuroimmunological review on cytokines involved in antidepressant treatment response. Hum Psychopharmacol 2010;25(3):201–15.

[102] Seymour PA, Schmidt AW, Schulz DW. The pharmacology of CP-154,526, a non-peptide antagonist of the CRH1 receptor: a review. CNS Drug Rev 2003;9(1):57–96.

[103] Blazquez A, Mas S, Plana MT, Lafuente A, Lazaro L. Fluoxetine pharmacogenetics in child and adult populations. Eur Child Adolesc Psychiatry 2012;21(11):599–610.

[104] Papiol S, Arias B, Gasto C, Gutierrez B, Catalan R, Fananas L. Genetic variability at HPA axis in major depression and clinical response to antidepressant treatment. J Affect Disord 2007;104(1–3):83–90.

[105] Binder EB, Owens MJ, Liu W, Deveau TC, Rush AJ, Trivedi MH, et al. Association of polymorphisms in genes regulating the corticotropin-releasing factor system with antidepressant treatment response. Arch Gen Psychiatry 2010;67(4):369–79.

[106] Wong ML, Dong C, Maestre-Mesa J, Licinio J. Polymorphisms in inflammation-related genes are associated with susceptibility to major depression and antidepressant response. Mol Psychiatry 2008;13(8):800–12.

[107] van Rossum EF, Binder EB, Majer M, Koper JW, Ising M, Modell S, et al. Polymorphisms of the glucocorticoid receptor gene and major depression. Biol Psychiatry 2006;59(8):681–8.

[108] Panarelli M, Holloway CD, Fraser R, Connell JM, Ingram MC, Anderson NH, et al. Glucocorticoid receptor polymorphism, skin vasoconstriction, and other metabolic intermediate phenotypes in normal human subjects. J Clin Endocrinol Metab 1998;83(6):1846–52.

[109] Rosmond R, Chagnon YC, Holm G, Chagnon M, Perusse L, Lindell K, et al. A glucocorticoid receptor gene marker is associated with abdominal obesity, leptin, and dysregulation of the hypothalamic-pituitary-adrenal axis. Obes Res 2000;8(3):211–8.

[110] Ukkola O, Rosmond R, Tremblay A, Bouchard C. Glucocorticoid receptor Bcl I variant is associated with an increased atherogenic profile in response to long-term overfeeding. Atherosclerosis 2001;157(1):221–4.

[111] Binder EB, Salyakina D, Lichtner P, Wochnik GM, Ising M, Putz B, et al. Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. Nat Genet 2004;36(12):1319–25.

[112] Kirchheiner J, Lorch R, Lebedeva E, Seeringer A, Roots I, Sasse J, et al. Genetic variants in FKBP5 affecting response to antidepressant drug treatment. Pharmacogenomics 2008;9(7):841–6.

[113] Lekman M, Laje G, Charney D, Rush AJ, Wilson AF, Sorant AJ, et al. The FKBP5-gene in depression and treatment response–an association study in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Cohort. Biol Psychiatry 2008;63(12):1103–10.

[114] Zobel A, Schuhmacher A, Jessen F, Hofels S, von Widdern O, Metten M, et al. DNA sequence variants of the FKBP5 gene are associated with unipolar depression. Int J Neuropsychopharmacol 2010;13(5):649–60.

[115] Miller AH, Maletic V, Raison CL. Inflammation and its discontents: the role of cytokines in the pathophysiology of major depression. Biol Psychiatry 2009;65(9):732–41.

[116] Yu YW, Chen TJ, Hong CJ, Chen HM, Tsai SJ. Association study of the interleukin-1 beta (C-511T) genetic polymorphism with major depressive disorder, associated symptomatology, and antidepressant response. Neuropsychopharmacology 2003;28(6):1182–5.

[117] Tadic A, Rujescu D, Muller MJ, Kohnen R, Stassen HH, Szegedi A, et al. Association analysis between variants of the interleukin-1beta and the interleukin-1 receptor antagonist gene and antidepressant treatment response in major depression. Neuropsychiat Dis Treat 2008;4(1):269–76.

[118] Baune BT, Dannlowski U, Domschke K, Janssen DG, Jordan MA, Ohrmann P, et al. The interleukin 1 beta (IL1B) gene is associated with failure to achieve remission and impaired emotion processing in major depression. Biol Psychiatry 2010;67(6):543–9.

[119] Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W, et al. Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. Am J Psychiatry 2010;167(5):555–64.

[120] Chessell IP, Hatcher JP, Bountra C, Michel AD, Hughes JP, Green P, et al. Disruption of the P2X7 purinoceptor gene abolishes chronic inflammatory and neuropathic pain. Pain 2005;114(3):386–96.

[121] Lucae S, Salyakina D, Barden N, Harvey M, Gagne B, Labbe M, et al. P2RX7, a gene coding for a purinergic ligand-gated ion channel, is associated with major depressive disorder. Hum Mol Genet 2006;15(16):2438–45.

[122] Stein DJ, Daniels WM, Savitz J, Harvey BH. Brain-derived neurotrophic factor: the neurotrophin hypothesis of psychopathology. CNS Spectr 2008;13(11):945–9.

[123] Brunoni AR, Lopes M, Fregni F. A systematic review and meta-analysis of clinical studies on major depression and BDNF levels: implications for the role of neuroplasticity in depression. Int J Neuropsychopharmacol 2008;11(8):1169–80.

[124] Nibuya M, Morinobu S, Duman RS. Regulation of BDNF and trkB mRNA in rat brain by chronic electroconvulsive seizure and antidepressant drug treatments. J Neurosci 1995;15(11):7539–47.

[125] Egan MF, Kojima M, Callicott JH, Goldberg TE, Kolachana BS, Bertolino A, et al. The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. Cell 2003;112(2):257–69.

[126] Petryshen TL, Sabeti PC, Aldinger KA, Fry B, Fan JB, Schaffner SF, et al. Population genetic study of the brain-derived neurotrophic factor (BDNF) gene. Mol Psychiatry 2010;15(8):810–5.

[127] Warner-Schmidt JL, Duman RS. VEGF is an essential mediator of the neurogenic and behavioral actions of antidepressants. Proc Natl Acad Sci USA 2007;104(11):4647–52.

[128] Iga J, Ueno S, Yamauchi K, Numata S, Tayoshi-Shibuya S, Kinouchi S, et al. Gene expression and association analysis of vascular endothelial growth factor in major depressive disorder. Prog Neuropsychopharmacol Biol Psychiatry 2007;31(3):658–63.

[129] Tsai SJ, Hong CJ, Liou YJ, Chen TJ, Chen ML, Hou SJ, et al. Haplotype analysis of single nucleotide polymorphisms in the vascular endothelial growth factor (VEGFA) gene and antidepressant

treatment response in major depressive disorder. Psychiatry Res 2009;169(2):113–7.

[130] Talbot K, Eidem WL, Tinsley CL, Benson MA, Thompson EW, Smith RJ, et al. Dysbindin-1 is reduced in intrinsic, glutamatergic terminals of the hippocampal formation in schizophrenia. J Clin Invest 2004;113(9):1353–63.

[131] Chen G, Hasanat KA, Bebchuk JM, Moore GJ, Glitz D, Manji HK. Regulation of signal transduction pathways and gene expression by mood stabilizers and antidepressants. Psychosom Med 1999;61(5):599–617.

[132] Ruiz-Velasco V, Ikeda SR. A splice variant of the G protein beta 3-subunit implicated in disease states does not modulate ion channels. Physiol Genomics 2003;13(2):85–95.

[133] Hong CJ, Chen TJ, Yu YW, Tsai SJ. Response to fluoxetine and serotonin 1A receptor (C-1019G) polymorphism in Taiwan Chinese major depressive disorder. Pharmacogenomics J 2006;6(1):27–33.

[134] Kang RH, Hahn SW, Choi MJ, Lee MS. Relationship between G-protein beta-3 subunit C825T polymorphism and mirtazapine responses in Korean patients with major depression. Neuropsychobiology 2007;56(1):1–5.

[135] Kato M, Wakeno M, Okugawa G, Fukuda T, Takekita Y, Hosoi Y, et al. Antidepressant response and intolerance to SSRI is not influenced by G-protein beta3 subunit gene C825T polymorphism in Japanese major depressive patients. Prog Neuropsychopharmacol Biol Psychiatry 2008;32(4):1041–4.

[136] Grimes CA, Jope RS. The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. Prog Neurobiol 2001;65(4):391–426.

[137] Beaulieu JM, Gainetdinov RR, Caron MG. Akt/GSK3 signaling in the action of psychotropic drugs. Annu Rev Pharmacol Toxicol 2009;49:327–47.

[138] Beaulieu JM, Zhang X, Rodriguiz RM, Sotnikova TD, Cools MJ, Wetsel WC, et al. Role of GSK3 beta in behavioral abnormalities induced by serotonin deficiency. Proc Natl Acad Sci USA 2008;105(4):1333–8.

[139] Kwok JB, Hallupp M, Loy CT, Chan DK, Woo J, Mellick GD, et al. GSK3B polymorphisms alter transcription and splicing in Parkinson's disease. Ann Neurol 2005;58(6):829–39.

[140] Benedetti F, Dallaspezia S, Lorenzi C, Pirovano A, Radaelli D, Locatelli C, et al. Gene-gene interaction of glycogen synthase kinase 3-beta and serotonin transporter on human antidepressant response to sleep deprivation. J Affect Disord 2012;136(3):514–9.

[141] Adli M, Hollinde DL, Stamm T, Wiethoff K, Tsahuridu M, Kirchheiner J, et al. Response to lithium augmentation in depression is associated with the glycogen synthase kinase 3-beta -50T/C single nucleotide polymorphism. Biol Psychiatry 2007;62(11):1295–302.

[142] Tsai SJ, Liou YJ, Hong CJ, Yu YW, Chen TJ. Glycogen synthase kinase-3beta gene is associated with antidepressant treatment response in Chinese major depressive disorder. Pharmacogenomics J 2008;8(6):384–90.

[143] Dwivedi Y, Rao JS, Rizavi HS, Kotowski J, Conley RR, Roberts RC, et al. Abnormal expression and functional characteristics of cyclic adenosine monophosphate response element binding protein in postmortem brain of suicide subjects. Arch Gen Psychiatry 2003;60(3):273–82.

[144] Blendy JA. The role of CREB in depression and antidepressant treatment. Biol Psychiatry 2006;59(12):1144–50.

[145] Serretti A, Chiesa A, Calati R, Massat I, Linotte S, Kasper S, et al. A preliminary investigation of the influence of CREB1 gene on treatment resistance in major depression. J Affect Disord 2011;128(1–2):56–63.

[146] O'Brien FE, Dinan TG, Griffin BT, Cryan JF. Interactions between antidepressants and P-glycoprotein at the blood-brain barrier: clinical significance of *in vitro* and *in vivo* findings. Brit J Pharmacol 2012;165(2):289–312.

[147] Eichelbaum M, Fromm MF, Schwab M. Clinical aspects of the MDR1 (ABCB1) gene polymorphism. Ther Drug Monit 2004;26(2):180–5.

[148] de Klerk OL, Nolte IM, Bet PM, Bosker FJ, Snieder H, den Boer JA, et al. ABCB1 gene variants influence tolerance to selective serotonin reuptake inhibitors in a large sample of Dutch cases with major depressive disorder. Pharmacogenomics J 2013;13(4):349–53.

[149] Noordam R, Aarts N, Hofman A, van Schaik RH, Stricker BH, Visser LE. Association between genetic variation in the ABCB1 gene and switching, discontinuation, and dosage of antidepressant therapy: results from the Rotterdam Study. J Clin Psychopharmacol 2013;33(4):546–50.

[150] Singh AB, Bousman CA, Ng CH, Byron K, Berk M. ABCB1 polymorphism predicts escitalopram dose needed for remission in major depression. Transl Psychiatry 2012;2:e198.

[151] Porcelli S, Fabbri C, Spina E, Serretti A, De Ronchi D. Genetic polymorphisms of cytochrome P450 enzymes and antidepressant metabolism. Expert Opin Drug Metab Toxicol 2011;7(9):1101–15.

[152] Kirchheiner J, Brosen K, Dahl ML, Gram LF, Kasper S, Roots I, et al. CYP2D6 and CYP2C19 genotype-based dose recommendations for antidepressants: a first step towards subpopulation-specific dosages. Acta Psychiatr Scand 2001;104(3):173–92.

[153] Laje G, McMahon FJ. Genome-wide association studies of antidepressant outcome: a brief review. Prog Neuropsychopharmacol Biol Psychiatry 2010;35(7):1553–7.

[154] Lin PI, Vance JM, Pericak-Vance MA, Martin ER. No gene is an island: the flip-flop phenomenon. Am J Hum Genet 2007;80(3):531–8.

[155] Shen X, Carlborg O. Beware of risk for increased false positive rates in genome-wide association studies for phenotypic variability. Front Genet 2013;4:93.

[156] Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009;460(7256):748–52.

[157] Broer L, Lill CM, Schuur M, Amin N, Roehr JT, Bertram L, et al. Distinguishing true from false positives in genomic studies: p values. Eur J Epidemiol 2013;28(2):131–8.

[158] Lin WY, Lee WC. Improving power of genome-wide association studies with weighted false discovery rate control and prioritized subset analysis. PLoS One 2012;7(4):e33716.

[159] Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, et al. A genomewide association study of citalopram response in major depressive disorder. Biol Psychiatry 2010;67(2):133–8.

[160] Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, et al. A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. Arch Gen Psychiatry 2009;66(9):966–75.

[161] Crumbley C, Wang Y, Kojetin DJ, Burris TP. Characterization of the core mammalian clock component, NPAS2, as a REV-ERBalpha/RORalpha target gene. J Biol Chem 2010;285(46):35386–92.

[162] Hunter AM, Leuchter AF, Power RA, Muthen B, McGrath PJ, Lewis CM, et al. A genome-wide association study of a sustained pattern of antidepressant response. J Psychiatr Res 2013;47(9):1157–65.

[163] Tammiste A, Jiang T, Fischer K, Magi R, Krjutskov K, Pettai K, et al. Whole-exome sequencing identifies a polymorphism in the BMP5 gene associated with SSRI treatment response in major depression. J Psychopharmacol 2013;27(10):915–20.

[164] Reines A, Bernier LP, McAdam R, Belkaid W, Shan W, Koch AW, et al. N-cadherin prodomain processing regulates synaptogenesis. J Neurosci 2012;32(18):6323–34.

[165] Galtrey CM, Fawcett JW. The role of chondroitin sulfate proteoglycans in regeneration and plasticity in the central nervous system. Brain Res Rev 2007;54(1):1–18.

[166] Uher R. Genes, environment, and individual differences in responding to treatment for depression. Harv Rev Psychiatry 2011;19(3):109–24.

[167] Antypa N, Drago A, Serretti A. Genomewide interaction and enrichment analysis on antidepressant response. Psychol Med 2013:1–13.

[168] Tansey KE, Guipponi M, Perroud N, Bondolfi G, Domenici E, Evans D, et al. Genetic predictors of response to serotonergic and noradrenergic antidepressants in major depressive disorder: a genome-wide analysis of individual-level data and a meta-analysis. PLoS Medicine 2012;9(10):e1001326.

[169] GENDP Investigators, MARS Investigators, STAR*D Investigators. Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. Am J Psychiatry 2013;170(2):207–17.

[170] Clark SL, Adkins DE, Aberg K, Hettema JM, McClay JL, Souza RP, et al. Pharmacogenomic study of side-effects for antidepressant treatment options in STAR*D. Psychol Med 2012;42(6):1151–62.

[171] Adkins DE, Clark SL, Aberg K, Hettema JM, Bukszar J, McClay JL, et al. Genome-wide pharmacogenomic study of citalopram-induced side effects in STAR*D. Transl Psychiatry 2012;2:e129.

[172] Holmans P. Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. Adv Genet 2010;72:141–79.

[173] Serretti A, Fabbri C. Shared genetics among major psychiatric disorders. Lancet 2013;381(9875):1339–41.

[174] Gotlib IH, Hamilton JP. Bringing genetics back to psychiatric endophenotypes. Biol Psychiatry 2012;71(1):2–3.

[175] Mamdani F, Lopez J, Berlim M, Labbe A, Turecki G. Transcriptomic and epigenetic correlates of antidepressant response. In: XIXth world congress of psychiatric genetics: genes to biology; 2011. p. 8.

[176] Serretti A, Olgiati P, Bajo E, Bigelli M, De Ronchi D. A model to incorporate genetic testing (5-HTTLPR) in pharmacological treatment of major depressive disorders. World J Biol Psychiatry 2011;12(7):501–15.

[177] Olgiati P, Bajo E, Bigelli M, De Ronchi D, Serretti A. Should pharmacogenetics be incorporated in major depression treatment? Economic evaluation in high- and middle-income European countries. Prog Neuropsychopharmacol Biol Psychiatry 2012;36(1):147–54.

[178] Nutt DJ, Baldwin DS, Clayton AH, Elgie R, Lecrubier Y, Montejo AL, et al. Consensus statement and research needs: the role of dopamine and norepinephrine in depression and antidepressant treatment. J Clin Psychiatry 2006;67(Suppl 6):46–9.

Chapter 27

Personalized Medicine of Alzheimer's Disease

Ramón Cacabelos, Pablo Cacabelos, and Clara Torrellas
Chair of Genomic Medicine, Camilo José Cela University, Madrid, Spain; EuroEspes Biomedical Research Center, Corunna, Spain

Chapter Outline

27.1 Overview 563
27.2 Toward a Personalized Medicine for Dementia and Neurodegenerative Disorders 564
27.3 Genomics of Alzheimer's Disease 565
27.4 Pathogenic Events 565
  27.4.1 Genomic Defects 566
  27.4.2 Epigenetic Phenomena 567
  27.4.3 Cerebrovascular Dysfunction 567
  27.4.4 Phenotypic Expression of Amyloid Deposits and Neurofibrillary Tangles 567
  27.4.5 Neuronal Apoptosis 567
  27.4.6 Neurotransmitter Deficits 568
  27.4.7 Oxidative Stress 568
  27.4.8 Cholesterol and Lipid Metabolism Dysfunction 568
  27.4.9 Neuroinflammation and Immunopathology 568
  27.4.10 Neurotoxic Factors 568
  27.4.11 Other Players 568
27.5 Biomarkers and Comorbidity 569
  27.5.1 Neuropathology 569
  27.5.2 Structural and Functional Neuroimaging 569
  27.5.3 Neurophysiology 569
  27.5.4 Biochemistry of Body Fluids 570
  27.5.5 Genomics and Proteomics 570
27.6 Therapeutic Strategies 571
  27.6.1 Immunotherapy 571
27.7 Pharmacogenomics 573
  27.7.1 Pathogenic Genes 573
  27.7.2 Genes Involved in the Mechanism of Action of CNS Drugs 574
  27.7.3 Genes Involved in Drug Metabolism 574
  27.7.4 Drug Transporters 580
  27.7.5 Pleiotropic Activity of ***APOE*** in Dementia 580
  27.7.6 Pharmacogenomics of Antidementia Drugs 581
  27.7.7 Multifactorial Therapy 582
27.8 Future Perspective 584
27.9 Conclusion 585
Appendix A 586
Appendix B 608
References 611

27.1 OVERVIEW

Since the identification of its pathogenic features by Alois Alzheimer in 1906, more than 90,000 papers have been published on Alzheimer's disease (AD) to date (2.5 million references on cancer since 1818; 1.6 million on cardiovascular disorders since 1927; and 1.01 million on central nervous system disorders since 1893) [1]. The number of people affected by dementia is becoming a public and socioeconomic concern in many countries all over the world, independent of economic conditions. The growth of the elderly population is a common phenomenon in both developed and developing countries, bringing about future challenges in terms of health policy and disability rates.

In the United States, rates for the leading causes of death are heart disease (200.2 per 100,000), cancer (180.7 per 100,000), and stroke (43.6 per 100,000). AD is the fifth leading cause of death in people older than 65 years of age, representing 71,600 deaths per year. AD affects approximately 5.4 million individuals in the United States and is estimated to affect up to 16 million by 2050 [2]. Disability caused by senility and dementia affects 9.2 per 1000 in the population aged 65–74 years, 33.5 per 1000 in those within the 75–84 range, and 83.4 per 1000 in the population over 85 years

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00027-X
© 2014 Elsevier Inc. All rights reserved.

[3,4]. In low- to middle-income countries, dementia makes the largest contribution to disability, with a median population-attributable prevalence fraction of 25.1%, followed by stroke (11.4%), limb impairment (10.5%), arthritis (9.9%), depression (8.3%), eyesight problems (6.8%), and gastrointestinal impairments (6.5%) [5].

In Western countries, AD is the most prevalent form of dementia (45–60%), followed by vascular dementia (30–40%), and mixed dementia (10–20%), which in people older than 85 years of age may account for more than 80% of cases.

The different forms of dementia pose several challenges to society and to the scientific community: (1) they represent an epidemiological problem and a socioeconomic, psychological, and family burden; (2) most of them have an obscure/complex pathogenesis; (3) their diagnosis is not easy and lacks specific biomarkers; and (4) their treatment is difficult and inefficient.

In terms of economic burden, approximately 10–20% of direct costs are associated with pharmacological treatment, with a gradual increase that parallels the severity of the disease. A Canadian study [6] shows that the mean total cost to treat patients with very mild AD is $367 per month, compared with $4063 per month for patients with severe or very severe AD. Only 20–30% of patients with dementia respond appropriately to conventional drugs, and the onset of adverse drug reactions imposes the need for other drugs to neutralize side effects, thus multiplying the initial cost of the pharmacological treatment and the health risk for the patients [7]. Wimo et al. [8] studied the economic impact of dementia in Europe in the EU-funded Eurocode project and found that the total cost of dementia in EU27 countries in 2008 was estimated to be €160 billion (€22,000 per dementia patient per year), of which 56% were costs of informal care. The corresponding costs for the whole of Europe were €177 billion. Informal caregiver costs were the largest cost component, accounting for about half to just over 60% of total societal costs, depending on the country and AD severity [9].

In addition (and related) to the problem of direct and indirect costs for the management of dementia, there is an alarming abuse of inappropriate psychotropic drug consumption worldwide. Antipsychotic medications are taken by more than 30% of elderly patients with dementia [10], and conventional antipsychotics are associated with a higher risk of all-cause mortality among nursing home residents [11].

Abuse, misuse, self-prescription, and uncontrolled medical prescription of CNS drugs are becoming major problems with unpredictable consequences for brain health. The pharmacological management of dementia is an issue of special concern because of the polymedication required to modulate its symptomatic complexity where cognitive decline, behavioral changes, and psychomotor deterioration coexist. In parallel, a growing body of fresh knowledge is emerging on the pathogenesis of dementia, together with data on the neurogenomics and pharmacogenomics of CNS disorders. The incorporation of this new armamentarium of molecular pathology and genomic medicine into daily medical practice, together with educational programs for the correct use of drugs, must help researchers and clinicians to (1) understand AD pathogenesis; (2) establish an early diagnosis; and (3) optimize therapeutics either as a preventive strategy or as formal symptomatic treatment [7,12].

27.2 TOWARD A PERSONALIZED MEDICINE FOR DEMENTIA AND NEURODEGENERATIVE DISORDERS

Common features of neurodegenerative disorders include the following:

- Polygenic/complex disorders in which genetic, epigenetic, and environmental factors are involved
- Deterioration of higher activities of the CNS
- Multifactorial dysfunction in several brain circuits
- Accumulation of toxic proteins in the nervous tissue

For instance, the neuropathological hallmarks of AD (amyloid deposition in senile plaques, neurofibrillary tangle formation, and neuronal loss) are merely the phenotypic expression of a pathogenic process in which different gene clusters and their products are potentially involved [7,12].

A large number of the genes that form the structural architecture of the human genome are expressed in the brain in a time-dependent manner along the lifespan. The cellular complexity of the CNS (10^3 different cell types) and synapses (each of the 10^{11} neurons in the brain having around 10^3–10^4 synapses with a complex multiprotein structure integrated by 10^3 different proteins) requires very powerful technology for gene expression profiling, which is still in its very early stages and is not devoid of technical obstacles and limitations [13]. Transcripts of 16,896 genes have been measured in different CNS regions. Each region possesses its own unique transcriptome fingerprint that is independent of age, gender, and energy intake. Fewer than 10% of genes are affected by age, diet, or gender, with most of these changes occurring between middle and old age. Gender and energy restriction have robust influences on the hippocampal transcriptome of middle-aged animals. Prominent functional groups of age- and energy-sensitive genes are those encoding proteins involved in DNA damage responses, mitochondrial and proteasome functions, cell fate determination, and synaptic vesicle trafficking [14].

The introduction of novel procedures in an integral genomic medicine protocol for CNS disorders and dementia is imperative in drug development and in clinical practice in order to improve diagnostic accuracy and to optimize therapeutics. Personalized strategies, adapted to the complexity of each case, are essential to depict a clinical profile based on specific biomarkers correlating with individual genomic profiles [7,15].

Our understanding of the pathophysiology of CNS disorders and dementia has advanced dramatically during the last 30 years, especially in terms of their molecular pathogenesis and genetics. The drug treatment of CNS disorders has also made remarkable strides with the introduction of many new drugs for the treatment of schizophrenia, depression, anxiety, epilepsy, Parkinson's disease, and AD, among many other quantitatively and qualitatively important neuropsychiatric disorders.

Improvement in terms of clinical outcome, however, has fallen short of expectations, with up to one-third of patients continuing to experience clinical relapse or unacceptable medication-related side effects in spite of efforts to identify optimal treatment regimes with one or more drugs. Potential reasons for this historical setback might be: (1) that the molecular pathology of most CNS disorders is still poorly understood; (2) that drug targets are inappropriate, not fitting into the real etiology of the disease; (3) that most treatments are symptomatic but not antipathogenic; (4) that the genetic component of most CNS disorders is poorly defined; and (5) that the understanding of genome–drug interactions is very limited [7,12].

The optimization of CNS therapeutics requires the establishment of new postulates regarding (1) the costs of medicines, (2) the assessment of protocols for multifactorial treatment in chronic disorders, (3) the implementation of novel therapeutics addressing causative factors, and (4) the establishment of pharmacogenomic strategies for drug development [12]. Personalized therapeutics based on individual genomic profiles implies the characterization of five types of gene clusters:

- Genes associated with disease pathogenesis
- Genes associated with the mechanism of action of drugs
- Genes associated with drug metabolism (phase I and II reactions)
- Genes associated with drug transporters
- Pleiotropic genes involved in multifaceted cascades and metabolic reactions

27.3 GENOMICS OF ALZHEIMER'S DISEASE

More than 3000 genes distributed across the human genome have been screened for association with AD during the past 30 years [16]. In the Alzgene database [17] there are 695 genes potentially associated with AD, of which the top ten are (in decreasing order of importance): *APOE* (19q13.2), *BIN1* (2q14), *CLU* (8p21–p12), *ABCA7* (19p13.3), *CR1* (1q32), *PICALM* (11q14), *MS4A6A* (11q12.1), *CD33* (19q13.3), *MS4A4E* (11q12.2), and *CD2AP* (6p12). Potentially defective genes associated with AD represent about 1.39% (35,252.69 Kb) of the human genome, which is integrated by 36,505 genes (3,095,677.41 Kb). The highest number of AD-related defective genes concentrate on chromosomes 10 (5.41%; 7337.83 Kb), 21 (4.76%; 2289.15 Kb), 7 (1.62%; 2584.26 Kb), 2 (1.56%; 3799.67 Kb), 19 (1.45%; 854.54 Kb), 9 (1.42%; 2010.62 Kb), 15 (1.23%; 1264.4 Kb), 17 (1.19%; 970.16 Kb), 12 (1.17%; 1559.9 Kb), and 6 (1.15%; 1968.22 Kb), with the highest proportion (related to the total number of genes mapped on a single chromosome) located on chromosome 10 and the lowest on chromosome Y [18] (Figure 27.1).

The genetic and epigenetic defects identified in AD can be classified into four major categories: Mendelian mutations; susceptibility SNP; mtDNA mutations; and epigenetic changes. Mendelian mutations affect genes directly linked to AD, including 32 mutations in the amyloid beta precursor protein (*APP*) gene (21q21)(AD1), 165 mutations in the presenilin 1 (*PSEN1*) gene (14q24.3)(AD3), and 12 mutations in the presenilin 2 (*PSEN2*) gene (1q31–q42) (AD4) [16–20]. *PSEN1* and *PSEN2* are important determinants of γ-secretase activity responsible for proteolytic cleavage of APP and NOTCH receptor proteins. Mendelian mutations are very rare in AD (1:1000). Mutations in exons 16 and 17 of the *APP* gene appear with a frequency of 0.30% and 0.78%, respectively, in AD patients. Likewise, *PSEN1*, *PSEN2*, and microtubule-associated protein Tau (*MAPT*)(17q21.1) mutations are present in less than 2% of cases. Mutations in these genes confer specific phenotypic profiles to patients with dementia: amyloidogeneic pathology associated with *APP*, *PSEN1*, and *PSEN2* mutations and tauopathy associated with *MAPT* mutations represent the two major pathogenic hypotheses for AD [16–21].

Multiple polymorphic risk variants can increase neuronal vulnerability to premature death (see Appendix A). Among these susceptibility genes, the apolipoprotein E (*APOE*) gene (19q13.2)(AD2) is the most prevalent as a risk factor for AD, especially in those subjects harboring the *APOE-4* allele (Figure 27.2), whereas carriers of the *APOE-2* allele might be protected against dementia. *APOE*-related pathogenic mechanisms are also associated with brain aging and with the neuropathological hallmarks of AD [16].

27.4 PATHOGENIC EVENTS

The dual amyloidogenic-tauopathic theory of AD has dominated the pathogenic universe of AD-related neurodegeneration (and divided the research community) for the past 50 years, nourished by the presence of *APP*, *PSEN1*, *PSEN2*, and *MAPT* mutations in a very small number of cases with early-onset AD. Nevertheless, this theory does not explain AD pathogenesis in full, and consequently novel (or complementary) theories have been emerging recently and during the past decades. A summary of the pathogenic events in AD is given in the following sections.

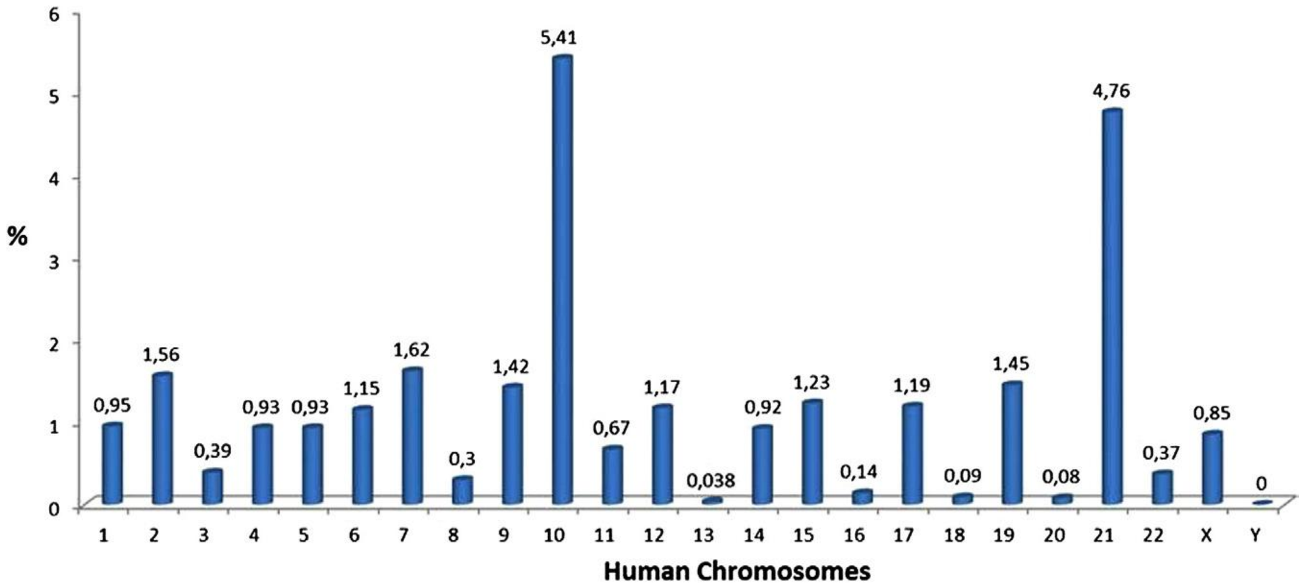
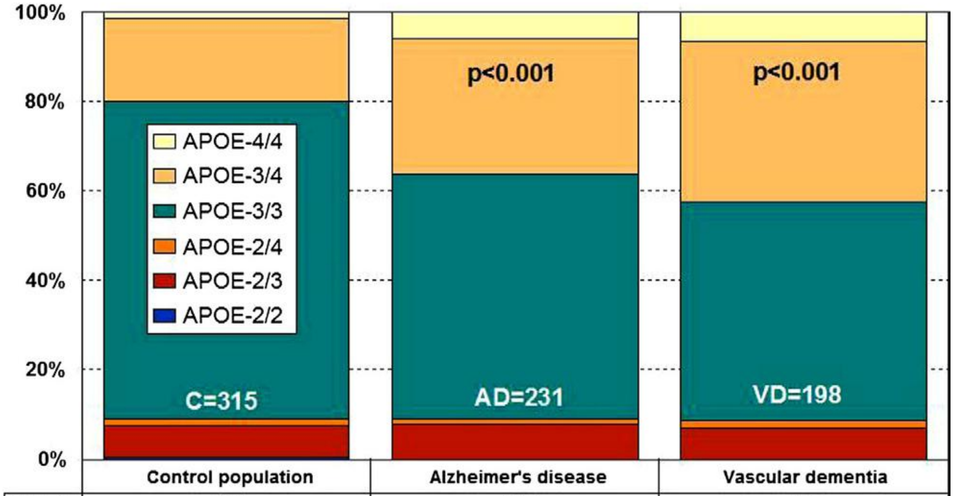


FIGURE 27.1 Distribution of AD-related genes in the human genome.



	Control population	Alzheimer's disease	Vascular dementia
APOE-4/4	1,59	6,06	6,57
APOE-3/4	18,41	30,3	35,85
APOE-3/3	71,11	54,55	48,99
APOE-2/4	1,27	1,3	1,52
APOE-2/3	7,3	7,79	7,07
APOE-2/2	0,32	0	0

FIGURE 27.2 Distribution and frequency of *APOE* genotypes in AD and vascular dementia. *Source: Adapted from Cacabelos* [18].

27.4.1 Genomic Defects

As a complex polygenic/multifactorial disorder, in which hundreds of polymorphic variants of risk might be involved (Appendix A, Figure 27.1), AD fulfils the "golden rule" of complex disorders, according to which the larger the number of genetic defects distributed in the human genome, the earlier the onset of the disease and the poorer its therapeutic response to conventional treatments; conversely, the smaller the number of pathogenic SNPs, the later the onset of the disease and the better its therapeutic response to different pharmacological interventions [12,16,22–28]. Genetic variation associated with different diseases interferes with microRNA-mediated regulation by creating, destroying, or modifying microRNA (miRNA) binding sites. miRNA-target variability is a ubiquitous phenomenon in the adult human brain which may influence gene expression in physiological and pathological conditions. AD-related SNPs interfere with miRNA gene regulation and affect AD susceptibility. Significant interactions include target SNPs present in seven genes related to AD prognosis with the miRNAs miR-214, -23a and -23b, -486-3p, -30e*, -143, -128, -27a and -27b, -324-5p, and -422a. The dysregulated miRNA network contributes to aberrant gene expression in AD [29–31].

27.4.2 Epigenetic Phenomena

Epigenetic factors have emerged as important mediators of development and aging, gene–gene and gene–environmental interactions, and the pathophysiology of complex disorders. Major epigenetic mechanisms (DNA methylation, histone modifications and chromatin remodeling, and noncoding RNA regulation) may contribute to AD pathology [30,31].

27.4.3 Cerebrovascular Dysfunction

Vascular and metabolic dysfunctions are key components in AD pathology throughout the course of disease. Although common denominators between vascular and metabolic dysfunction are oxidative stress and Aβ [32], genetic factors and cardiovascular risk factors may also account for the cerebrovascular damage present in AD [33]. Inherited polymorphisms of the vascular susceptibility gene Ninjurin2 (*NINJ2*) are associated with AD risk [34]. Endothelial dysfunction has been implicated as a crucial event in the development of AD.

Breakdown of the blood–brain barrier (BBB) as a result of disruption of tight junctions and transporters leads to increased leukocyte transmigration and is an early event in the pathology of many CNS disorders. BBB breakdown leads to neuroinflammation and oxidative stress, with mitochondrial dysfunction. The high concentration of mitochondria in cerebrovascular endothelial cells might account for the sensitivity of the BBB to oxidant stressors [35,36].

Chronic brain hypoperfusion may be sufficient to induce premature neuronal death and dementia in vulnerable subjects [16,23–25,37–39]. APOE-related changes in cortical oxygenation and hemoglobin consumption are evident, as revealed by brain optical topography analysis, and reflect that *APOE-4* carriers exhibit deficient brain hemodynamics and a poorer panneocortical oxygenation than do *APOE-3* or *APOE-2* carriers [18]. Hypoperfusion in frontal, parietal, and temporal regions is a common finding in AD. White matter hyperintensities (WMH) correlate with age and with disease severity [40].

Cerebral amyloid angiopathy (CAA) accounts for the majority of primary lobal intracerebral hemorrhages (ICH) among the elderly, and represents the cause of 20% of spontaneous ICHs in patients over 70 years of age. The basis for this disease process is the deposition and formation of eventually destructive amyloid plaques in the walls of brain vessels, predominantly arterial but not excluding venules and capillaries. CAA and CAA-associated microhemorrhages may also participate in the pathogenesis of AD [41]. Aβ deposition in asymptomatic elderly individuals is associated with lobar MH (LMH).

LMH is present in 30.8% of AD, 35.7% of MCI, and 19.1% of controls [42]. Neurovascular dysfunction in AD leads to reduced clearance across the BBB and accumulation of neurotoxic Aβ peptides in the brain. The ABC transport protein P-glycoprotein (P-gp, ABCB1) is involved in the export of Aβ from the brain into the blood. *P-gp*, *LRP1*, and *RAGE* mRNA expression is reduced in mice treated with $A\beta_{1–42}$. In addition to the age-related decrease in P-gp expression, $A\beta_{1–42}$ itself downregulates the expression of P-gp and other Aβ transporters, which could exacerbate the intracerebral accumulation of Aβ and thereby accelerate neurodegeneration in AD and cerebral β-amyloid angiopathy [43].

27.4.4 Phenotypic Expression of Amyloid Deposits and Neurofibrillary Tangles

β-Amyloid deposits in senile and neuritic plaques and hyperphosphorylated tau proteins in neurofibrillary tangles (NFT) are extracellular and intracellular expressions, respectively, of the AD neuropathological phenotype, together with selective neuronal loss in hippocampal and neocortical regions. Aβ plaque in the brain is the primary (postmortem) diagnostic criterion of AD. The main component of senile plaques is Aβ, a 39–43 amino acid peptide, generated by the proteolytic cleavage of amyloid precursor protein (APP) by the action of beta- and gamma-secretases. Aβ is neurotoxic, and this neurotoxicity is related to its aggregation state [16–21].

27.4.5 Neuronal Apoptosis

Neuronal loss is a pathognomonic finding in AD and the final common path of multiple pathogenic mechanisms leading to neurodegeneration in dementia. Atrophy of the medial temporal lobe, especially the hippocampus and the parahippocampal gyrus, is considered to be AD's most predictive structural brain biomarker. The medial and posterior parts of the parietal lobe seem to be preferentially affected, compared to the other parietal lobe parts [18].

27.4.6 Neurotransmitter Deficits

An imbalance of different neurotransmitters (glutamate, acetylcholine, noradrenaline, dopamine, serotonin, and some neuropeptides) has been proposed as the neurobiological basis of behavioral symptoms in AD. Altered reuptake of neurotransmitters by vesicular glutamate transporters (VGLUTs), excitatory amino acid transporters (EAATs), the vesicular acetylcholine transporter (VAChT), the serotonin reuptake transporter (SERT), or the dopamine reuptake transporter (DAT) are involved in the neurotransmission imbalance in AD. Protein and mRNA levels of VGLUTs, EAAT1-3, VAChT, and SERT are reduced in the disease [44].

27.4.7 Oxidative Stress

Oxidative damage is a classic pathogenic mechanism of neurodegeneration [36,45]. It is greater in brain tissue from patients with AD than age-matched controls. Tayler et al. [46] studied the timing of this damage in relation to other pathogenic AD processes. Antioxidant capacity is elevated in AD and directly related to disease severity as indicated by the Braak tangle stage and the amount of insoluble Aβ. Accumulation of Aβ has been shown in brain mitochondria of AD patients and in AD transgenic mouse models. The presence of Aβ in mitochondria leads to free radical generation and neuronal stress.

A novel mitochondrial Aβ-degrading enzyme, presequence protease (Pre), has been identified in the mitochondrial matrix. hPreP activity is decreased in AD human brains and in the mitochondrial matrix of AD transgenic mouse brains (TgmAβPP and TgmAβPP/ABAD). Mitochondrial fractions isolated from AD brains and TgmAβPP mice have higher levels of 4-hydroxynonenal, an oxidative product. Cytochrome c oxidase activity is significantly reduced in the AD mitochondria. Decreased PreP proteolytic activity, possibly due to enhanced ROS production, may contribute to Aβ accumulation in mitochondria, leading to mitochondrial toxicity and neuronal death in AD [47].

27.4.8 Cholesterol and Lipid Metabolism Dysfunction

Cholesterol seems to be intimately linked with the generation of amyloid plaques, which are central to AD pathogenesis. APOE variants are determinants in cholesterol metabolism and diverse forms of dyslipoproteinemia [12,48]. Cholesterol protects the Aβ-induced neuronal membrane disruption and inhibits beta-sheet formation of Aβ on the lipid bilayer [49]. Jones et al. [50] found a significant over-representation of association signals in pathways related to cholesterol metabolism and the immune response in both of the two largest genome-wide association studies for LOAD.

27.4.9 Neuroinflammation and Immunopathology

Several genes associated with immune regulation and inflammation show polymorphic variants of risk in AD, and abnormal levels of diverse cytokins have been reported in the brain, CSF, and plasma of AD patients [16,23]. The activation of inflammatory cascades has been consistently demonstrated in AD pathophysiology, in which reactive microglia are associated with Aβ deposits and clearance. Resident microglia fail to trigger an effective phagocytic response to clear Aβ deposits, although they mainly exist in an "activated" state. Oligomeric Aβ (oAβ) can induce more potent neurotoxicity when compared with fibrillar Aβ (fAβ). $A\beta_{(1-42)}$ fibrils, not $A\beta_{(1-42)}$ oligomers, increase microglial phagocytosis [51]. Among several putative neuroinflammatory mechanisms, the TNF-α signaling system has a central role in this process. In AD, TNF-α levels are altered in serum and CSF. The abnormal production of inflammatory factors may accompany the progression from mild cognitive impairment (MCI) to dementia. Abnormal activation of the TNF-α signaling system, represented by increased expression of sTNFR1, is associated with a higher risk of progression from MCI to AD [52].

27.4.10 Neurotoxic Factors

Old and new theories suggest that different toxic agents, from metals (e.g., aluminium, copper, zinc, iron) to biotoxins and pesticides, might contribute to neurodegeneration. Dysfunctional homeostasis of transition metals is believed to play a role in AD pathogenesis [18].

27.4.11 Other Players

Many novel pathogenic mechanisms potentially involved in AD neurodegeneration have been proposed in recent times. Moreover, there has been a revival of some old hypotheses. Examples of pathogenic players in AD, other than those just discussed, include the Ca^{2+} hypothesis, insulin resistance, NGF imbalance, glycogen synthase kinase-3 (GSK-3), advanced glycation end products (AGEs) and their receptors (RAGE), the efflux transporter P-glycoprotein (P-gp), c-Abl tyrosine kinase, post-transcriptional protein alterations that compromise the proteasome system and the chaperone machinery (HSPB8–BAG3), autophagy as a novel Aβ-generating pathway, hypocretin (orexin), cathepsin B, Nogo receptor proteins, adipocytokines and $CD34^{+}$ progenitor cells, CD147, impairment of synaptic plasticity (PSD-95), anomalies in neuronal cell division and apoptosis, stem cell factor (SCF), telomere shortening, deficiency in repair of nuclear and mitochondrial DNA damage, and microDNAs [18].

27.5 BIOMARKERS AND COMORBIDITY

AD's phenotypic features represent the biomarkers to be used as diagnostic predictors and the expression of pathogenic events to be modified with an effective therapeutic intervention. Important differences have been found in the AD population (as compared with healthy subjects) in different biological parameters, including blood pressure, glucose, cholesterol and triglyceride levels, transaminase activity, hematological parameters, metabolic factors, thyroid function, brain hemodynamic parameters, and brain mapping activity [7,23–25,53–59].

These clinical differences are clear signs of comorbidity rather than typical features of AD. Blood pressure values, glucose levels, and cholesterol levels are higher in AD than in healthy elderly subjects. Approximately 20% of AD patients are hypertensive, 25% are diabetics, 50% are hypercholesterolemic, and 23% are hypertriglyceridemic. More than 25% of patients exhibit high GGT activity, 5–10% show anemic conditions, 30–50% show an abnormal cerebrovascular function characterized by poor brain perfusion, and more than 60% have an abnormal electroencephalographic pattern, especially in frontal, temporal, and parietal regions, as revealed by quantitative EEG (qEEG) or computerized mapping [7,12,23,54]. Significant differences are currently seen between females and males, indicating the effect of gender on the phenotypic expression of the disease. In fact, the prevalence of dementia is 10–15% higher in females than in males from 65–85 years of age. All of these parameters are highly relevant when treating AD patients, because some of them reflect a concomitant pathology that also needs therapeutic consideration.

AD biomarkers can be differentiated into several categories: (1) neuropathological markers; (2) structural and functional neuroimaging markers; (3) neurophysiological markers (EEG, qEEG, brain mapping); (4) biochemical markers in body fluids (e.g., blood, urine, saliva, CSF); and (5) genomic markers (structural and functional genomics, proteomics, metabolomics).

27.5.1 Neuropathology

Plaques and tangles in the hippocampus and cortex are still considered the seminal findings in AD neuropathology and are the conventional means of establishing the boundary between amyloidopathies and tauopathies; however, both phenotypic markers are also present in normal brains, in more than 60% of cases with traumatic brain injury, and in many other brain disorders [60].

27.5.2 Structural and Functional Neuroimaging

Structural and functional neuroimaging techniques (MRI, fMRI, PET, SPECT) are essential tools in the diagnosis of dementia, although the specificity of visual observations in degenerative forms of dementia is of doubtful value. Nevertheless, these procedures are irreplaceable for a differential diagnosis. There is a characteristic regional impairment in AD that involves mainly the temporo–parietal association cortices, the mesial temporal structures, and, to a more variable degree, the frontal association cortex. This pattern of functional impairment can provide a biomarker for diagnosis of AD and other neurodegenerative dementias at the clinical stage of mild cognitive impairment, and for monitoring its progression. Healthy young *APOE ε4* carriers have smaller hippocampal volumes than *APOE ε2* carriers.

The difference in hippocampal morphology is cognitively/clinically silent in young adulthood, but can render *APOE ε4* carriers more prone to the later development of AD, possibly because of lower reserve cognitive capacity [61]. LOAD patients exhibit a selective parahippocampal white matter (WM) loss, while EOAD patients experience a more widespread pattern of posterior WM atrophy. The distinct regional distribution of WM atrophy reflects the topography of gray matter (GM) loss. *ApoE ε4* status is associated with a greater parahippocampal WM loss in AD. The greater WM atrophy in EOAD than in LOAD fits with the evidence that EOAD is a more aggressive form of the disease [62]. FDG-PET is quantitatively more accurate than perfusion SPECT.

Regional metabolic and blood flow changes are closely related to clinical symptoms, and most areas involved in these changes also develop significant cortical atrophy. FDG-PET is complementary to amyloid PET, which targets a molecular marker that does not have a close relation to current symptoms. FDG-PET is expected to play an increasing role in diagnosing patients at an early stage of AD and in clinical trials of drugs aimed at preventing or delaying the onset of dementia [63]. Functional neuroimaging biomarkers are becoming popular, with the introduction of novel tracers for brain amyloid deposits. Amyloid deposition causes severe damage to neurons many years before onset of dementia via a cascade of several downstream effects.

Positron emission tomography (PET) tracers for amyloid plaque are desirable for early diagnosis of AD, particularly to enable preventative treatment once effective therapeutics is available. The amyloid imaging tracers flutemetamol, florbetapir, and florbetaben labeled with ^{18}F have been developed for PET. These tracers are currently undergoing formal clinical trials to establish whether they can be used to accurately image fibrillary amyloid, and to distinguish patients with AD from normal controls and those with other diseases that cause dementia [63].

27.5.3 Neurophysiology

There is a renewed interest in the use of computerized brain mapping as a diagnostic aid and as a monitoring tool in AD [64]. Electroencephalography (EEG) studies in AD show an attenuation of average power within the alpha band (7.5–13 Hz) and an increase in power in the theta band (4–7 Hz) [65]. *APOE* genotypes influence brain bioelectrical activity in AD. In general, *APOE-4* carriers tend to exhibit a slower EEG pattern from early stages [16,18,66].

27.5.4 Biochemistry of Body Fluids

Other biomarkers of potential interest include cerebrospinal fluid (CSF) and peripheral levels of $A\beta_{42}$, protein tau, histamine, interleukins, and some other novel candidate markers such as chitinase 3-like 1 (CHI3L1) protein [7,16,25,67–69]. The concentration of the 42-amino-acid form of Aβ ($A\beta_{1-42}$) is reduced in the CSF of AD patients, which is believed to reflect the AD pathology, with plaques in the brain acting as sinks. Novel C-truncated forms of Aβ ($A\beta_{1-14}$, $A\beta_{1-15}$, and $A\beta_{1-16}$) were identified in human CSF. The presence of these small peptides is consistent with a catabolic amyloid precursor protein cleavage pathway by β- followed by α-secretase. $A\beta_{1-14}$, $A\beta_{1-15}$, and $A\beta_{1-16}$ increase dose-dependently in response to γ-secretase inhibitor treatment, while $A\beta_{1-42}$ levels are unchanged [70].

Kester et al. [71] investigated change over time in CSF levels of amyloid-beta 40 and 42 ($A\beta_{40}$ and $A\beta_{42}$), total tau (tau), tau phosphorylated at threonine 181 (ptau-181), isoprostane, neurofilaments heavy (NfH) and neurofilaments light (NfL). $A\beta_{42}$, tau, and tau phosphorylated at threonine 181 differentiated between diagnosis groups, whereas isoprostane, NfH, and NfL did not. In contrast, effects of follow-up time were found only for nonspecific CSF biomarkers: levels of NfL decreased, and levels of isoprostane, $A\beta_{40}$, and tau increased over time. An increase in isoprostane was associated with progression of mild cognitive impairment in AD and with cognitive decline. Contrary to AD-specific markers, nonspecific CSF biomarkers show change over time, which potentially can be used to monitor disease progression in AD.

27.5.5 Genomics and Proteomics

Structural markers are represented by SNPs in genes associated with AD, polygenic cluster analysis, and genome-wide studies (GWSs). Functional markers attempt to correlate genetic defects with specific phenotypes (genotype–phenotype correlations). In proteomic studies, several candidate CSF protein biomarkers have been assessed in neuropathologically confirmed AD, nondemented (ND) elderly controls, and non-AD dementias (NADD). Markers selected included apolipoprotein A-1 (ApoA1), hemopexin (HPX), transthyretin (TTR), pigment epithelium-derived factor (PEDF), $A\beta_{1-40}$, $A\beta_{1-42}$, total tau, phosphorylated tau, α-1 acid glycoprotein (A1GP), haptoglobin, zinc α-2 glycoprotein (Z2GP), and apolipoprotein E (ApoE). Concentrations of $A\beta_{1-42}$, ApoA1, A1GP, ApoE, HPX, and Z2GP differed significantly among AD, ND, and NADD subjects. The CSF concentrations of these three markers distinguished AD from ND subjects with 84% sensitivity and 72% specificity, with 78% of subjects correctly classified.

By comparison, $A\beta_{1-42}$ alone gave 79% sensitivity and 61% specificity, with 68% of subjects correctly classified. For the diagnostic discrimination of AD from NADD, only the concentration of $A\beta_{1-42}$ was significantly related to diagnosis, with a sensitivity of 58% and a specificity of 86% [72]. Carrying the *APOE-ε4* allele was associated with a significant decrease in CSF $A\beta_{1-42}$ concentrations in middle-aged and older subjects. In AD, $A\beta_{1-42}$ levels are significantly lower in *APOEε4* carriers compared to noncarriers. These findings demonstrate significant age effects on CSF $A\beta_{1-42}$ and pTau181 across the lifespan, and also suggest that a decrease in $A\beta_{1-42}$, but an increase in pTau181 CSF levels, is accelerated by the *APOEε4* genotype in middle-aged and older adults with normal cognition [73].

Han et al. [74] carried out a GWAS to better define the genetic backgrounds of normal cognition, mild cognitive impairment (MCI), and AD in terms of changes in CSF levels of $A\beta_{1-42}$, T-tau, and P-tau181P. CSF $A\beta_{1-42}$ levels decreased with *APOE* gene dose for each subject group. T-tau levels tended to be higher among AD cases than among normal subjects. *CYP19A1* "aromatase" (rs2899472), *NCAM2*, and multiple SNPs located on chromosome 10 near the *ARL5B* gene demonstrated the strongest associations with $A\beta_{1-42}$ in normal subjects.

Two genes found to be near the top SNPs, *CYP19A1* (rs2899472) and *NCAM2* (rs1022442), have been reported as genetic factors related to the progression of AD. In AD subjects, *APOE ε2/ε3* and *ε2/ε4* genotypes were associated with elevated T-tau levels, and the *ε4/ε4* genotype was associated with elevated T-tau and P-tau181P levels. Blood-based markers reflecting core pathological features of AD in presymptomatic individuals are likely to accelerate the development of disease-modifying treatments.

Thambisetty et al. [75] performed a proteomic analysis to discover plasma proteins associated with brain Aβ burden in nondemented older individuals. A panel of 18 2DGE plasma protein spots effectively discriminated between individuals with high and low brain Aβ. Mass spectrometry identified these proteins, many of which have established roles in Aβ clearance, including a strong signal from ApoE. A strong association was observed between plasma ApoE concentration, and Aβ burden in the medial temporal lobe. Targeted voxel-based analysis localized this association to the hippocampus and entorhinal cortex. *APOE ε4* carriers also showed greater Aβ levels in several brain regions relative to *ε4* noncarriers. Both peripheral concentration of the ApoE protein and the *APOE* genotype may be related to early neuropathological changes in brain regions vulnerable to AD pathology even in the nondemented elderly.

27.6 THERAPEUTIC STRATEGIES

Modern therapeutic strategies in AD are aimed at interfering with the main pathogenic mechanisms potentially involved in AD [7,12,16,18,23,24,28,53–59] (Box 27.1). Starting in the early 1990s, the neuropharmacology of AD was dominated by acetylcholinesterase inhibitors, represented by tacrine, donepezil, rivastigmine, and galantamine [76–78]. Memantine, a partial NMDA antagonist, was introduced in the 2000s for the treatment of severe dementia [79]; and the first clinical trials with immunotherapy, to reduce amyloid burden in senile plaques, were withdrawn due to severe ADRs [80,81]. After the initial promise of β- and γ-secretase inhibitors [82,83] and novel vaccines [84,85] devoid of severe side effects, during the past few years no relevant drug candidates have dazzled the scientific community with their capacity to halt disease progression; however, a large number of novel therapeutic strategies for the pharmacological treatment of AD have been postulated, with some apparent effects in preclinical studies (see Box 27.1).

27.6.1 Immunotherapy

There are two main modalities of immunotherapy for AD: (1) passive immunotherapy, with the administration of monoclonal Aβ-specific antibodies [86]; and (2) active immunization with the $A\beta_{42}$ antigen [87,88] or Aβ-conjugated synthetic fragments bound to a carrier protein, thus avoiding potential problems associated with mounting a T-cell response directly against Aβ [89]. A new approach—delivering $A\beta_{42}$ in a novel immunogen-adjuvant manner consisting of sphingosine-1-phosphate (S1P)-containing liposomes, administered to APP/PS1 transgenic mice before and after the detection of AD-like pathology in the brain—has recently been developed [85].

The results from this novel vaccine (EB101) indicate that active immunization significantly prevents and reverses the progression of AD-like pathology and also clears prototypical neuropathological hallmarks in transgenic mice. This new approach strongly induces T-cell, B-cell, and microglial immune response activation, avoiding the Th1 inflammatory reaction [90].

The rationale for amyloid immunotherapy in AD [91] is based on the following assumptions:

- β-amyloid plaques and their aggregated, proto-fibrillar, and oligomeric precursors contain immunologic neo-epitopes that are absent from the full-length amyloid precursor protein (APP), as well as from its soluble proteolytic derivatives restricted to brain tissue; consequently, β-amyloid-based immunotherapies designed to selectively target pathologic neo-epitopes present on Aβ oligomers, protofibrils, or fibrils should not cause autoimmune disease in unaffected tissues throughout the organism
- β-amyloid buildup precedes neurodegeneration and functional loss, and either the prevention of its formation or its removal can be expected to result in the slowing or the prevention of neurodegeneration
- β-amyloid can cause the formation of neurofibrillary tangles in vivo and in vitro. The removal of β-amyloid, or the prevention of its buildup, has the potential not only to correct β-amyloid-related toxicity, but also to prevent the formation of neurofibrillary tangles
- Conformational changes of endogenously occurring proteins and the formation of insoluble aggregates are commonly associated with neurodegeneration and brain disease, so the removal or prevention of these pathologic protein aggregates is also a therapeutic goal in the principle of immunotherapy
- Immunotherapy works in experimental animals and in initial clinical trials: both active immunization and passive antibody transfer consistently reduce brain β-amyloid load, improve β-amyloid-related memory impairments, and protect neurons against degeneration in many independent experiments using different mouse models and primates [90]

Since Aβ immunotherapy has a limited clearance effect of tau aggregates in dystrophic neurites, the development of an alternative therapy that directly targets pathological tau has become crucial. Increased levels of tau oligomers have been observed in the early stage of AD, prior to the detection of neurofibrillary tangles (NFT) formed by aggregation and accumulation of the microtubule-associated protein tau [92]. Several approaches have been taken to treat AD by targeting tau, such as the following:

1. The inhibition of tau hyperphosphorylation, by a kinase inhibitor of soluble aggregated tau formation, which also prevents related motor deficits [93].
2. Activation of the proteolytic pathway, by the degrading action of calpain [94] and puromycin-sensitive aminopeptidase [95].
3. The stabilization of microtubules, treating tauopathies by functionally binding and stabilizing microtubules with mt-binding protein tau [96] and paclitaxel, a drug proven effective in restoring affected axonal transport and motor impairments [97].
4. Tau clearance by immunotherapy in this case, the tau active vaccination uses phosphorylated antigens of tau fragments associated with neurofibrillary tangles [98] that results in an efficient reduction of both soluble and insoluble tau active fragments, reducing phosphorylated NFTs in AD-like mouse brains.

Preclinical studies have shown clear evidence that Aβ immunization therapy provides protection and reverses the pathological effects of AD in transgenic mouse models [99]. This strategy seems to improve cognition performance [100]

BOX 27.1 Experimental Strategies for the Pharmacological Treatment of Alzheimer's Disease

- New cholinesterase inhibitors
- Cholinergic receptor agonists
- Monoamine regulators
- Diverse natural compounds derived from vegetal sources:
  - Alkaloids from the calabar bean (*Physostigma venenosum*)
  - Huperzine A from *Huperzia serrata*
  - Galantamine from the snowdrop *Galanthus woronowii*
  - Cannabinoids (cannabidiol) from *Cannabis sativa*
  - Saffron (*Crocus sativus*)
  - Ginseng (*Panax species*)
  - Sage (*Salvia species*)
  - Lemon balm (*Melissa officinalis*)
  - *Polygala tenuifolia*
  - Nicotine from *Nicotiana species*
  - Grape seed polyphenolic extracts
  - Fuzhisan, a Chinese herbal medicine
  - Resveratrol
  - Xanthoceraside
  - Garlic (*Allium sativum*)
  - Linarin from *Mentha arvensis* and *Buddleja davidii*
  - Carotenoids (e.g., retinoic acid, all-trans retinoic acid, lycopene and β-carotene)
  - Curcumin from the rhizome of *Curcuma longa*
  - Decursinol from the roots of *Angelica gigas*
  - Bacopa monniera LINN *(Syn. Brahmi)*
  - Olive oil
  - Phytoestrogens
  - Walnut extract
  - *Erigeron annuus* leaf extracts
  - Epigallocatechin-3-gallate
  - Luteolin
  - The brown algae (*Ecklonia cava*)
  - Gami-Chunghyuldan (standardized multiherbal medicinal formula)
  - *Punica granatum* extracts
- Plants of different origin:
  - Yizhi Jiannao
  - Drumstick tree (*Moringa oleifera*)
  - Ginkgo/Maidenhair tree (*Ginkgo biloba*)
  - Sicklepod (*Cassia obtisufolia*)
  - Sal Leaved Desmodium (*Desmodium gangeticum*)
  - Lemon Balm (*Melissa officinalis*)
  - Garden sage, common sage (*Salvia officinalis*)
- Immunotherapy and treatment options for tauopathies:
  - Tau kinase inhibitors
  - 2-Aminothiazoles
  - Phosphoprotein phosphatase 2A (PP2A) inhibitors
  - c-Jun N-terminal kinase (JNKs) inhibitors
  - p38 MAP kinase inhibitors (CNI-1493)
  - Harmine (β-carboline alkaloid)
- Immunotherapy and Aβ breakers for AD-related amyloidopathy:
  - Active and passive immunization
- Secretase inhibitors (β- and γ-)
- Neostatins
- Neurosteroids
- Phosphodiesterase inhibitors
- Protein phosphatase methylesterase-1 inhibitors
- Histone deacetylase inhibitors
- mTOR inhibitors
- Peroxisome proliferator-activated receptor agonists
- P-glycoprotein regulators
- Nuclear receptor agonists
- Glycogen synthase kinase-3β (GSK-3β) regulators
- Histamine H3 receptor inverse agonists
- Estrogens
- Kynurenine 3-monooxygenase inhibitors
- Chaperones (small heat shock proteins (sHSPs); Hsp90 inhibitors and HSP inducers)
- microRNAs (miRNAs) and gene silencing (RNA interference)(RNAi)
- Miscellaneous strategies:
  - Sodium fullerenolate
  - Glucagon-like peptide -1 (GLP-1)
  - Chemokines
  - Macrophage inflammatory protein-2 (MIP-2)
  - Stromal cell-derived factor-1α (SDF-1α)
  - Cyclooxygenase-1 and cyclooxygenase-2 inhibitors
  - Bone morphogenetic protein 9 (BMP-9)
  - Granulocyte colony stimulating factor (G-CSF)/AMD3100 (CXCR4 antagonist)
  - Vitamins (A, B, C, D)
  - ω-3 Polyunsaturated fatty acids (n-3 PUFAs)
  - Docosahexaenoic acid (DHA, C22:6 n-3)
  - Sphingosylphosphorylcholine
  - Citidine-5-diphosphocholine (CDP-choline)
  - Cathepsin B inhibitors
  - Pituitary adenylate cyclase–activating polypeptide
  - NAP (Davunetide)
  - Transcription factor specificity protein 1 (Sp1) inhibitors (tolfenamic acid)
  - TNF inhibitors:
    - 2-(2,6-Dioxopiperidin-3-yl)phthalimidine EM-12 dithiocarbamates
    - N-substituted 3-(Phthalimidinp-2-yl)-2,6-dioxopiperidines
    - 3-substituted 2,6-Dioxopiperidines
  - Pyrrolo[3,2-e][1,2,4]triazolo[1,5-a]pyrimidine (SEN1176)
  - Latrepirdine
  - Leucettines
  - Dihydropyridines (inhibitors of L-type calcium channels)
  - Brain-penetrating angiotensin-converting enzyme (ACE) inhibitors
  - NADPH oxidase inhibitors (Apocynin)
  - Heterocyclic indazole derivatives (inhibitors of serum- and glucocorticoid-inducible-kinase 1 [*SGK1*])
  - IgG-single-chain Fv fusion proteins

after $A\beta_{42}$ immunization, in addition to causing an effective reduction in Aβ pathology. A recent immunization study has proven that a fragment of the Aβ peptide bound to polylysines activates the immune response that diminishes AD-like pathology in APP transgenic mice. This result reinforces the notion that the immune-conjugate approach is an effective means of Aβ immunotherapy, and also that the entire Aβ peptide is not necessary for its efficacy. It is in accordance with the hypothesis that specific antibodies directed against the amino-terminal and/or central region of the amyloid peptide provide beneficial protection against amyloid pathology. Passive immunization studies have also been conducted with promising experimental results, showing that a humoral response alone, without Aβ cellular response, is sufficient to reduce the β-amyloid burden and reverse memory deficits [101].

Among the drugs and vaccines currently under development to treat the pathological effects of AD, the most promising are bapineuzumab, solanezumab, CAD106, and EB101. Solanezumab is a monoclonal antibody raised against $A\beta_{13-28}$ that recognizes an epitope in the core of the amyloid peptide, binding selectively to soluble Aβ and with low affinity for the fibrillar Aβ form [102]. Thus, it presents fewer adverse events than does bapineuzumab, which binds to Aβ amyloid plaques more strongly than soluble Aβ [103]. There are a few other monoclonal antibodies against Aβ that have properties different from those of bapineuzumab, such as PF-04360365, which specifically targets the free carboxy-terminus of $A\beta_{1-40}$, MABT5102A, which binds with equally high affinity to Aβ monomers, oligomers, and fibrils, and GSK933776A, which targets the N-terminus of Aβ.

Specific anti-Aβ antibodies are present in pooled preparations of intravenous immunoglobulin (IVIg or IGIV), which has already been approved by the FDA for the treatment of a variety of neurological conditions. Current results from these studies have shown that IVIg treatment may also be an efficacious alternative approach in the treatment of AD neuropathologies [90,104].

Avoiding both the strong Th1 effects of the QS-21 adjuvant and the T-cell epitopes at the C-terminus of Aβ, CAD106 consists of a short N-terminal fragment of Aβ attached to a virus-like particle, with no additional adjuvant [105]. This therapeutic agent is currently in phase II trials. Affiris is testing two short 6-amino-peptides (AD01, AD02), administered with aluminum hydroxide as adjuvant, that mimic the free N-terminus of Aβ and therefore cause cross-reactivity with the native peptide in phase I trials [106]. In terms of prevention and therapeutic treatment, the EB101 vaccine showed for the first time the effectiveness of combining a liposomal immunogen-adjuvant with an Aβ antigen to induce an effective immunological response combined with an anti-inflammatory effect in preclinical studies using APP/PS1 transgenic mice [85,90].

The EB101 vaccine immunization process has shown a marked positive effect as a preventive and therapeutic treatment, reducing amyloidosis-induced inflammation as an effective Th2 immunomodulator. Moreover, this vaccine proved to stimulate innate immunity and enable effective phagocytosis to clear amyloid and neurofibrillary tangles, which are among the major hallmarks of AD-like neuropathology observed. A few other vaccines are currently under development, and recent studies have opened up new perspectives in the immunization approach to AD pathology; in particular, gene-gun-mediated genetic immunization with the $A\beta_{42}$ gene [107] shows that self-tolerance can be broken in order to produce a humoral response to the $A\beta_{42}$ peptide with minimal cellular response.

27.7 PHARMACOGENOMICS

AD patients may take 6–12 different drugs per day for the treatment of dementia-related symptoms, including memory decline (conventional antidementia drugs, neuroprotectants), behavioral changes (antidepressants, neuroleptics, sedatives, hypnotics), and functional decline. Such drugs may also be taken for the treatment of concomitant pathologies (epilepsy, cardiovascular and cerebrovascular disorders, parkinsonism, hypertension, dyslipidemia, anemia, arthrosis, etc). The co-administration of several drugs may cause side effects and ADRs in more than 60% of AD patients, who in 2–10% of cases require hospitalization. In more than 20% of patients, behavioral deterioration and psychomotor function can be severely altered by polypharmacy. The principal causes of these iatrogenic effects are (1) the inappropriate combination of drugs, and (2) the genomic background of the patient, which is responsible for his/her pharmacogenomic outcome.

Pharmacogenomics account for 30–90% of the variability in pharmacokinetics and pharmacodynamics. The genes involved in the pharmacogenomic response to drugs in AD fall into five major categories:

- Genes associated with AD pathogenesis and neurodegeneration (*APP*, *PSEN1*, *PSEN2*, *MAPT*, *PRNP*, *APOE*, and others)
- Genes associated with the mechanism of action of drugs (enzymes, receptors, transmitters, messengers)
- Genes associated with drug metabolism (phase I (*CYP*s) and phase II reactions (*UGT*s, *NAT*s))
- Genes associated with drug transporters (*ABC*s, *SLC*s)
- Pleiotropic genes involved in multifaceted cascades and metabolic reactions (*APOs*, *ILs*, *MTHFR*, *ACE*, *AGT*, *NOS*, etc) [18] (Figure 27.1)

27.7.1 Pathogenic Genes

In more than 100 clinical trials for dementia, *APOE* has been used as the only gene of reference for the pharmacogenomics of AD [7,12,15,16,22–28,53–59]. Several studies indicate that the presence of the *APOE-4* allele differentially affects the quality and extent of drug responsiveness in AD patients treated with cholinergic enhancers (tacrine, donepezil, galantamine, rivastigmine), neuroprotective compounds (nootropics), endogenous nucleotides (CDP-choline), immunotrophins (anapsos), neurotrophic factors (cerebrolysin), rosiglitazone, or combination therapies [108–110]; however, controversial results are frequently found that are due to methodological problems, study design, and patient recruitment in clinical trials.

The major conclusion in most studies is that *APOE-4* carriers are the worst responders to conventional treatments [7,12,15,16,22–28,53–59]. When *APOE* and *CYP2D6* genotypes are integrated in bigenic clusters and the *APOE+CYP2D6*-related therapeutic response to a combination therapy is analyzed in AD patients, it becomes clear that the presence of the *APOE-4/4* genotype is able to convert pure *CYP2D6*1/*1* extensive metabolizers (EMs) into full poor responders to conventional treatments, indicating the existence of a powerful influence of the *APOE-4* homozygous genotype on the drug-metabolizing capacity of pure *CYP2D6* EMs. In addition, a clear accumulation of *APOE-4/4* genotypes is observed among *CYP2D6* poor (PMs) and ultrarapid metabolizers (UMs) [12].

27.7.2 Genes Involved in the Mechanism of Action of CNS Drugs

Most genes associated with the mechanism of action of CNS drugs encode receptors, enzymes, and neurotransmitters on which psychotropic drugs act as ligands (agonists, antagonists), enzyme modulators (substrates, inhibitors, inducers), or neurotransmitter regulators (releasers, reuptake inhibitors) [111]. In the case of conventional antidementia drugs, tacrine, donepezil, rivastigmine and galantamine are cholinesterase inhibitors, and memantine is a partial NMDA antagonist (Table 27.1).

27.7.3 Genes Involved in Drug Metabolism

Drug metabolism includes phase I reactions (i.e., oxidation, reduction, hydrolysis) and phase II conjugation reactions (i.e., acetylation, glucuronidation, sulphation, methylation) (Table 27.2). The principal enzymes with polymorphic variants involved in phase I reactions are the following: cytochrome P450 monooxygenases (CYP3A4/5/7, CYP2E1, CYP2D6, CYP2C19, CYP2C9, CYP2C8, CYP2B6, CYP2A6, CYP1B1, CYP1A1/2), epoxide hydrolase, esterases, NQO1 (NADPH-quinone oxidoreductase), DPD (dihydropyrimidine dehydrogenase), ADH (alcohol dehydrogenase), and ALDH (aldehyde dehydrogenase). The major enzymes involved in phase II reactions include UGTs (uridine 5′-triphosphate glucuronosyl transferases), TPMT (thiopurine methyltransferase), COMT (catechol-O-methyltransferase), HMT (histamine methyl-transferase), STs (sulfotransferases), GST-A (glutathione S-transferase A), GST-P, GST-T, GST-M, NAT1 (N-acetyl transferase 1), NAT2, and others (Table 27.2).

Among these enzymes, CYP2D6, CYP2C9, CYP2C19, and CYP3A4/5 are the most relevant in the pharmacogenetics of CNS drugs [15,111] (Table 27.1). Approximately 18% of neuroleptics are major substrates of CYP1A2 enzymes, 40% of CYP2D6, and 23% of CYP3A4; 24% of antidepressants are major substrates of CYP1A2 enzymes, 5% of CYP2B6, 38% of CYP2C19, 85% of CYP2D6, and 38% of CYP3A4; 7% of benzodiazepines are major substrates of CYP2C19 enzymes, 20% of CYP2D6, and 95% of CYP3A4 [15,111]. Most CYP enzymes exhibit ontogenic-, age-, sex-, circadian-, and ethnic-related differences [112].

In dementia, as in any other CNS disorder, CYP genomics is a very important issue, since in practice more than 90% of patients with dementia are daily consumers of psychotropics. Furthermore, some acetylcholinesterase inhibitors (the most prescribed antidementia drugs worldwide) are metabolized via CYP enzymes (Table 27.1). Most CYP enzymes display highly significant ethnic differences, indicating that the enzymatic capacity of these proteins varies depending upon the polymorphic variants present in their coding CYP genes.

The practical consequence of this genetic variation is that the same drug can be differentially metabolized according to the genetic profile of each subject, and that, if an individual's pharmacogenomic profile is known, his/her pharmacodynamic response is potentially predictable. This is the cornerstone of pharmacogenetics. In this regard, the *CYP2D6*, *CYP2C19*, *CYP2C9*, and *CYP3A4/5* genes and their respective protein products deserve special consideration.

27.7.3.1 CYP2D6

CYP2D6 is a 4.38 kb gene with 9 exons mapped on 22q13.2. Four RNA transcripts of 1190–1684 bp are expressed in the brain, liver, spleen, and reproductive system, where 4 major proteins of 48–55 kDa (439–494 aa) are identified. It is a transport enzyme of the cytochrome P450 subfamily IID or multigenic cytochrome P450 superfamily of mixed-function monooxygenases. The cytochrome P450 proteins are monooxygenases which catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids, and other lipids. CYP2D6 localizes to the endoplasmic reticulum and is known to metabolize as many as 25% of commonly prescribed drugs, and more than 60% of current psychotropics. Its substrates include debrisoquine, an adrenergic-blocking drug; sparteine and propafenone, both antiarrhythmic drugs; and amitryptiline, an antidepressant. CYP2D6 is highly polymorphic in the population.

TABLE 27.1 Pharmacogenomic Profile of Antidementia Drugs

Donepezil	
Category	**Antidementia Agent/Cholinesterase Inhibitor**
Mechanism	Centrally active, reversible acetylcholinesterase inhibitor; increases acetylcholine available for synaptic transmission in CNS
Genes	
Pathogenic	APOE, CHAT
Mechanistic	CHAT, ACHE, BCHE
Metabolism: substrate	CYP2D6 (major), CYP3A4 (major), UGTs, ACHE
Metabolism: inhibitor	ACHE, BCHE
Transporter	ABCB1
Galantamine	
Category	**Antidementia Agent/Cholinesterase Inhibitor**
Mechanism	Reversible and competitive acetylcholinesterase inhibition leading to increased concentration of acetylcholine at cholinergic synapses; modulates nicotinic acetylcholine receptor; may increase glutamate and serotonin levels
Genes	
Mechanistic	APOE, APP
Pathogenic	ACHE, BCHE, CHRNA4, CHRNA7, CHRNB2
Metabolism: substrate	CYP2D6 (major), CYP3A4 (major), UGT1A1
Metabolism: inhibitor	ACHE, BCHE
Memantine	
Category	**Antidementia Drug; N-methyl-d-aspartate Receptor Antagonist**
Mechanism	Binds preferentially to NMDA receptor-operated cation channels; may act by blocking glutamate actions, mediated in part by NMDA receptors. Antagonists: GRIN2A, GRIN2B, GRIN3A, HTR3A, CHRFAM7A
Genes	
Pathogenic	APOE, PSEN1, MAPT
Mechanistic	GRIN2A, GRIN2B, GRIN3A, HTR3A, CHRFAM7A
Metabolism: inhibitor	CYP1A2 (weak), CYP2A6 (weak), CYP2B6 (strong), CYP2C9 (weak), CYP2C19 (weak), CYP2D6 (strong), CYP2E1 (weak), CYP3A4 (weak)
Pleiotropic	APOE, MAPT, MT-TK, PSEN1
Rivastigmine	
Category	**Antidementia Agent/Cholinesterase Inhibitor**
Mechanism	Increases acetylcholine in CNS through reversible inhibition of its hydrolysis by cholinesterase
Genes	
Pathogenic	APOE, APP, CHAT
Mechanistic	ACHE, BCHE, CHAT, CHRNA4, CHRNB2
Metabolism: inhibitor	ACHE, BCHE
Pleiotropic	APOE, MAPT

(Continued)

TABLE 27.1 Continued

Tacrine	
Category	**Antidementia agent/cholinesterase inhibitor**
Mechanism	Elevates acetylcholine in cerebral cortex by slowing degradation of acetylcholine
Genes	
Pathogenic	APOE
Mechanistic	ACHE, BCHE, CHRNA4, CHRNB2
Metabolism: substrate	CYP1A2 (major), CYP2D6 (minor), CYP3A4 (major)
Metabolism: inhibitor	ACHE, BCHE, CYP1A2 (weak)
Transporter	SCN1A
Pleiotropic	APOE, MTHFR, CES1, LEPR, GSTM1, GSTT1

Source: Cacabelos [113].

TABLE 27.2 Drug Metabolism-Related Genes

Phase I Enzymes	
Alcohol dehydrogenases	ADH1A, ADH1B, ADH1C, ADH4, ADH5, ADH6, ADH7, ADHFE1
Aldehyde dehydrogenases	ALDH1A1, ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, ALDH3A1, ALDH3A2, ALDH3B1, ALDH3B2, ALDH4A1, ALDH5A1, ALDH6A1, ALDH7A1, ALDH8A1, ALDH9A1, AOX1
Aldo-keto reductases	AKR1A1, AKR1B1, AKR1C1, AKR1D1
Amine oxidases	MAOA, MAOB, SMOX
Carbonyl reductases	CBR1, CBR3, CBR4
Cytidine deaminase	CDA
Cytochrome P450 family	CYP1A1, CYP1A2, CYP1B1, CYP2A6, CYP2A7, CYP2A13, CYP2B6, CYP2C18, CYP2C19, CYP2C8, CYP2C9, CYP2D6, CYP2D7P1, CYP2E1, CYP2F1, CYP2J2, CYP2R1, CYP2S1, CYP2W1, CYP3A4, CYP3A5, CYP3A7, CYP3A43, CYP4A11, CYP4A22, CYP4B1, CYP4F2, CYP4F3, CYP4F8, CYP4F11, CYP4F12, CYP4Z1, CYP7A1, CYP7B1, CYP8B1, CYP11A1, CYP11B1, CYP11B2, CYP17A1, CYP19A1, CYP20A1, CYP21A2, CYP24A1, CYP26A1, CYP26B1, CYP26C1, CYP27A1, CYP27B1, CYP39A1, CYP46A1, CYP51A1, POR, TBXAS1
Cytochrome b5 reductase	CYB5R3
Dihydropyrimidine dehydrogenase	DPYD
Esterases	AADAC, CEL, CES1, CES1P1, CES2, CES3, CES5A, ESD, GZMA, GZMB, PON1, PON2, PON3, UCHL1, UCHL3
Epoxidases	EPHX1, EPHX2
Flavin-containing monooxygenases	FMO1, FMO2, FMO3, FMO4, FMO5, FMO6P
Glutathione reductase/peroxidases	GSR, GPX1, GPX2, GPX3, GPX4, GPX5, GPX6, GPX7
Peptidases	DPEP1, METAP1
Prostaglandin-endoperoxide synthases	PTGS1, PTGS2
Short-chain dehydrogenases/reductases	DHRS1, DHRS2, DHRS3, DHRS4, DHRS7, DHRS9, DHRS12, DHRS13, DHRSX, HSD11B1, HSD17B10, HSD17B11, HSD17B14

TABLE 27.2 Continued

Superoxide dismutase	SOD1, SOD2
Xanthine dehydrogenase	XDH
	Phase II Enzymes
Amino acid transferases	AGXT, BAAT, CCBL1
Dehydrogenases	NQO1, NQO2, XDH
Esterases	CES1, CES2, CES3, CES4, CES5A
Glucuronosyl transferases	DDOST, UGT1A1, UGT1A10, UGT1A3, UGT1A4, UGT1A5, UGT1A6, UGT1A7, UGT1A8, UGT1A9, UGT2A1, UGT2A3, UGT2B10, UGT2B11, UGT2B15, UGT2B17, UGT2B28, UGT2B4, UGT2B7, UGT3A1, UGT8
Glutathione transferases	GSTA1, GSTA2, GSTA3, GSTA4, GSTA5, GSTCD, GSTK1, GSTM1, GSTM2, GSTM3, GSTM4, GSTM5, GSTO1, GSTO2, GSTP1, GSTT1, GSTT2, GSTZ1, MGST1, MGST2, MGST3, PTGES
Methyl transferases	AS3MT, ASMT, COMT, GAMT, GNMT, HNMT, INMT, NNMT, PNMT, TPMT
N-Acetyl transferases	AANAT, ACSL1, ACSL3, ACSL4, ACSM1, ACSM2B, ACSM3, GLYAT, NAT1, NAT2, NAA20, SAT1
Thioltransferase	GLRX
Sulfotransferases	SULT1A1, SULT1A2, SULT1A3, SULT1B1, SULT1C1, SULT1C2, SULT1C3, SULT1C4, SULT1E1, SULT2A1, SULT2B1, SULT4A1, SULT6B1, TST, CHST1, CHST2, CHST3, CHST4, CHST5, CHST6, CHST7, CHST8, CHST9, CHST10, CHST11, CHST12, CHST13, GAL3ST1

Note: See Appendix B for long-form names of genes listed.

There are 141 *CYP2D6* allelic variants, of which -100C>T, -1023C>T, -1659G>A, -1707delT, -1846G>A, -2549delA, -2613-2615delAGA, -2850C>T, -2988G>A, and -3183G>A represent the ten most important [113–115]. Different alleles result in the extensive, intermediate, poor, and ultrarapid metabolizer phenotypes, characterized by normal, intermediate, decreased, and multiplied ability to metabolize the enzyme's substrates, respectively. The hepatic cytochrome P450 system is responsible for the first phase in the metabolism and elimination of numerous endogenous and exogenous molecules and ingested chemicals. P450 enzymes convert these substances into electrophilic intermediates, which are then conjugated by phase II enzymes (e.g., UDP glucuronosyltransferases, N-acetyltransferases) to hydrophilic derivatives that can be excreted. According to the database of the *World Guide for Drug Use and Pharmacogenomics* [113], 982 drugs are *CYP2D6*-related: 371 are substrates, more than 300 are inhibitors, and 18 are CYP2D6 inducers.

In healthy subjects, extensive metabolizers (EMs) account for 55.71% of the population; intermediate metabolizers (IMs) account for 34.7%; poor metabolizers (PMs), 2.28%; and ultrarapid metabolizers (UMs), 7.31%. Remarkable worldwide interethnic differences exist in the frequency of the PM and UM phenotypes [116–118]. On average, approximately 6.28% of the world's population belongs to the PM category. Europeans (7.86%), Polynesians (7.27%), and Africans (6.73%) show the highest rate of PMs, whereas Orientals (0.94%) show the lowest [116]. The frequency of PMs among Middle Eastern populations, Asians, and Americans is in the range of 2–3%. *CYP2D6* gene duplications are relatively infrequent among Northern Europeans, but in East Africa the frequency of alleles with duplication of *CYP2D6* is as high as 29% [119]. In Europe, there is a North–South gradient in the frequency of PMs (6–12% of PMs in Southern European countries, and 2–3% of PMs in Northern latitudes) [111].

In AD, EMs, IMs, PMs, and UMs are 56.38%, 27.66%, 7.45%, and 8.51%, respectively, and in vascular dementia, they are, respectively, 52.81%, 34.83%, 6.74%, and 5.62% (Figure 27.3). There is an accumulation of AD-related risk genes in PMs and UMs. EMs and IMs are the best responders, and PMs and UMs are the worst responders to a combination therapy of cholinesterase inhibitors, neuroprotectants, and vasoactive substances. The pharmacogenetic response in AD appears to depend on the networking activity of genes involved in drug metabolism and genes involved in AD pathogenesis [7,12,15,16,22–28,53–59].

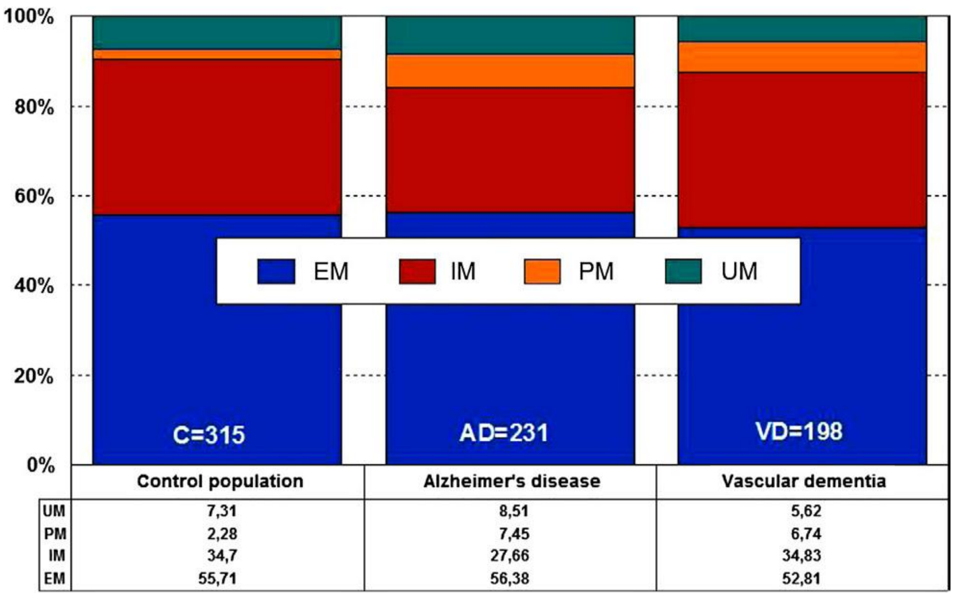


FIGURE 27.3 **Distribution and frequency of *CYP2D6* phenotypes in AD and vascular dementia.** EM—extensive metabolizer; IM—intermediate metabolizer; PM—poor metabolizer; UM—ultrarapid metabolizer. *Source: Adapted from Cacabelos* [18].

27.7.3.2 CYP2C9

CYP2C9 is a gene (50.71 kb) with 9 exons mapped on 10q24. An RNA transcript of 1860 bp is mainly expressed in hepatocytes, where a protein of 55.63 kDa (490 aa) can be identified. More than 600 drugs are *CYP2C9*-related: 311 act as substrates (177 major, 134 minor); 375, as inhibitors (92 weak, 181 moderate, and 102 strong); and 41 as inducers of the CYP2C9 enzyme [113]. There are 481 *CYP2C9* SNPs. By phenotype (Figure 27.4), in the control population, PMs represent 7.04%, IMs 32.39%, and EMs 60.56%. In AD, PMs, IMs, and EMs are 6.45%, 37.64%, and 55.91%, respectively, and in vascular dementia they are 3.61%, 28.92%, and 67.47%, respectively [18] (Figure 27.4).

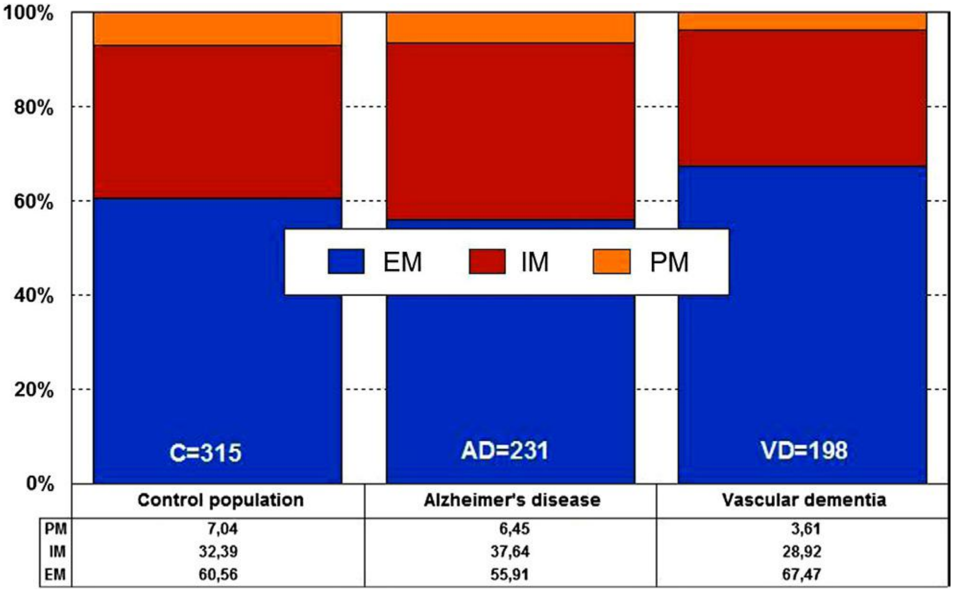
27.7.3.3 CYP2C19

CYP2C19 is a gene (90.21 kb) with 9 exons mapped on 10q24.1q24.3. RNA transcripts of 1901 bp, 2395 bp, and 1417 bp are expressed in liver cells, where a protein of 55.93 kDa (490 aa) has been identified. Nearly 500 drugs are *CYP2C19*-related, with 281 acting as substrates (151 major, 130 minor), 263 as inhibitors (72 weak, 127 moderate, and 64 strong), and 23 as inducers of the CYP2C19 enzyme [113]. About 541 SNPs have been detected in the *CYP2C19* gene. The frequencies of the three major *CYP2C19* geno-phenotypes in the control population are *CYP2C19-*1/*1*-EMs, 68.54%; *CYP2C19-*1/*2*-IMs, 30.05%; and *CYP2C19-*2/*2*-PMs, 1.41%. EMs, IMs, and PMs account for 69.89%, 30.11%, and 0%, respectively, in AD, and 66.27%, 30.12%, and 3.61%, respectively, in vascular dementia [18] (Figure 27.5).

27.7.3.4 CYP3A4/5

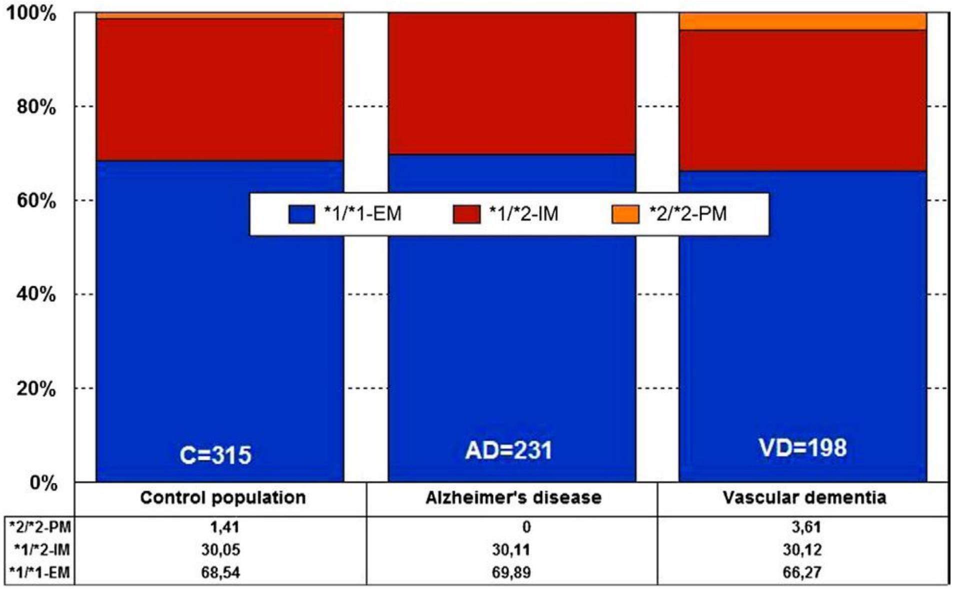
CYP3A4 is a gene (27.2 kb) with 13 exons mapped on 7q21.1. RNA transcripts of 2153 bp, 651 bp, 564 bp, 2318 bp, and 2519 bp are expressed in intestine, liver, prostate, and other tissues, where four protein variants of 57.34 kDa (503 aa), 17.29 kDa (153 aa), 40.39 kDa (353 aa), and 47.99 kDa (420 aa) have been identified. The human *CYP3A* locus contains the three *CYP3A* genes (*CYP3A4*, *CYP3A5,* and *CYP3A7*), three pseudogenes, and a novel *CYP3A* gene termed *CYP3A43*. The gene encodes a putative protein with 71.5–75.8% identity with the other CYP3A proteins. The predominant hepatic form is CYP3A4, but CYP3A5 contributes significantly to total liver CYP3A activity.

CYP3A4 metabolizes more than 1900 drugs: 1033 act as substrates (897 major, 136 minor); 696, as inhibitors (118 weak, 437 moderate, and 141 strong); and 241, as inducers of the CYP3A4 enzyme [113]. About 347 SNPs have been identified in the *CYP3A4* gene (*CYP3A4*1A*: wild-type), 25 of which are of clinical relevance. Concerning *CYP3A4/5* polymorphisms in AD, 82.75% of cases are EMs (*CYP3A5*3/*3*), 15.88% are IMs (*CYP3A5*1/*3*), and 1.37% are UMs (*CYP3A5*1/*1*). Unlike other human P450s (*CYP2D6*, *CYP2C19*), there is no evidence of a "null" allele for *CYP3A4* [113].



	Control population	Alzheimer's disease	Vascular dementia
PM	7,04	6,45	3,61
IM	32,39	37,64	28,92
EM	60,56	55,91	67,47

FIGURE 27.4 **Distribution and frequency of _CYP2C9_ phenotypes in AD and vascular dementia.** EM—extensive metabolizer; IM—intermediate metabolizer; PM—poor metabolizer. _Source: Adapted from Cacabelos_ [18].



	Control population	Alzheimer's disease	Vascular dementia
*2/*2-PM	1,41	0	3,61
*1/*2-IM	30,05	30,11	30,12
*1/*1-EM	68,54	69,89	66,27

FIGURE 27.5 **Distribution and frequency of _CYP2C19_ pheno- genotypes in AD and vascular dementia.** EM—extensive metabolizer; IM—intermediate metabolizer; PM—poor metabolizer. _Source: Adapted from Cacabelos_ [18].

27.7.3.5 CYP Clustering

The construction of a genetic map integrating the most prevalent _CYP2D6+CYP2C19+CYP2C9_ polymorphic variants in a trigenic cluster yields 82 different haplotype-like profiles. The most frequent trigenic genotypes in the AD population are _*1*1-*1*1-*1*1_ (25.70%), _*1*1-*1*2-*1*2_ (10.66%), _*1*1-*1*1-*1*1_ (10.45%), _*1*4-*1*1-*1*1_ (8.09%), _*1*4-*1*2-*1*1_ (4.91%), _*1*4-*1*1-*1*2_ (4.65%), and _*1*1-*1*3-*1*3_ (4.33%). These 82 trigenic genotypes represent 36 different pharmacogenetic phenotypes.

According to these trigenic clusters, only 26.51% of patients show a pure 3EM phenotype, 15.29% are 2EM1IM, 2.04% are pure 3IM, 0% are pure 3PM, and 0%

are 1UM2PM (the worst possible phenotype). This implies that only one-quarter of the population normally process the drugs that are metabolized via CYP2D6, CYP2C9, and CYP2C19 (approximately 60% of the drugs in current use) [12]. Taking into consideration the data available, it might be inferred that at least 20–30% of the AD population may exhibit an abnormal metabolism of cholinesterase inhibitors and/or other drugs that undergo oxidation via *CYP2D6*-related enzymes.

Approximately 50% of this population cluster shows an ultrarapid metabolism, requiring higher doses of cholinesterase inhibitors in order to reach a therapeutic threshold. The other 50% of the cluster exhibit a poor metabolism, displaying potential adverse events at low doses. If we take into account that approximately 60–70% of therapeutic outcomes depend on pharmacogenomic criteria (e.g., pathogenic mechanisms associated with AD-related genes), it can be postulated that pharmacogenetic and pharmacogenomic factors are responsible for 75–85% of therapeutic response (efficacy) in AD patients treated with conventional drugs [12,15–18,22–25,28,53–59].

27.7.4 Drug Transporters

ABC genes—especially *ABCB1* (ATP-binding cassette, subfamily B, member 1P-glycoprotein-1, P-gp1, Multidrug Resistance 1, MDR (17q21.12), *ABCC1* (9q31.1), *ABCG2* (White121q22.3), and other genes of this family—encode proteins that are essential for drug metabolism and transport. The multidrug efflux transporters P-gp, the multidrug resistance-associated protein 4 (MRP4), and the breast cancer resistance-protein (BCRP), located on endothelial cells lining the brain vasculature, play important roles in limiting the movement of substances into the brain and in enhancing their efflux from the brain.

Transporters also cooperate with phase I/phase II metabolism enzymes by eliminating drug metabolites. Their major features are their capacity to recognize drugs belonging to unrelated pharmacological classes and their redundancy, by which a single molecule can act as a substrate for different transporters. This ensures efficient neuroprotection against xenobiotic invasions. The pharmacological induction of ABC gene expression is a mechanism of drug interaction, which may affect substrates of the upregulated transporter; overexpression of MDR transporters confers resistance to anti-cancer agents and CNS drugs [120,121].

Also of importance for CNS pharmacogenomics are transporters encoded by genes of the solute carrier superfamily (SLC) and solute carrier organic (SLCO) transporter family, which are responsible for the transport of multiple endogenous and exogenous compounds, including folate (*SLC19A1*), urea (*SLC14A1*, *SLC14A2*), monoamines (*SLC29A4*, *SLC22A3*), aminoacids (*SLC1A5*, *SLC3A1*, *SLC7A3*, *SLC7A9*, *SLC38A1*, *SLC38A4*, *SLC38A5*, *SLC38A7*, *SLC43A2*, *SLC45A1*), nucleotides (*SLC29A2*, *SLC29A3*], fatty acids (*SLC27A1-6*), neurotransmitters (*SLC6A2*[noradrenaline transporter]), *SLC6A3*[dopamine transporter], *SLC6A4*[serotonin transporter, SERT], *SLC6A5*, *SLC6A6*, *SLC6A9*, *SLC6A11*, *SLC6A12*, *SLC6A14*, *SLC6A15*, *SLC6A16*, *SLC6A17*, *SLC6A18*, *SLC6A19*), glutamate (*SLC1A6*, *SLC1A7*), and others [122].

Some organic anion transporters (OAT), which belong to the solute carrier (SLC) 22A family, are also expressed at the BBB, and regulate the excretion of endogenous and exogenous organic anions and cations [123]. The transport of amino acids and di- and tripeptides is mediated by a number of different transporter families, and the bulk of oligopeptide transport is attributable to the activity of members of the *SLC15A* superfamily (peptide transporters 1 and 2 (*SLC15A1*[PepT1]) and *SLC15A2*[PepT2], and peptide/histidine transporters 1 and 2 (*SLC15A4*[PHT1] and *SLC15A3*[PHT2]). ABC and SLC transporters expressed at the BBB may cooperate to regulate the passage of different molecules into the brain [124]. Polymorphic variants in ABC and SLC genes may also be associated with pathogenic events in CNS disorders and drug-related safety and efficacy complications [111,122].

27.7.5 Pleiotropic Activity of *APOE* in Dementia

APOE is the prototypical paradigm of a pleiotropic gene with multifaceted activities in physiological and pathological conditions [16,22]. ApoE is consistently associated with the amyloid plaque marker for AD. *APOE*-4 may influence AD pathology interacting with APP metabolism and Aβ accumulation, enhancing hyperphosphorylation of tau protein and NFT formation, reducing choline acetyltransferase activity, increasing oxidative processes, modifying inflammation-related neuroimmunotrophic activity and glial activation, altering lipid metabolism, lipid transport, and membrane biosynthesis in sprouting and synaptic remodeling, and inducing neuronal apoptosis [16,23–25].

To address the complex misfolding and aggregation that initiates the toxic cascade resulting in AD, Petrlova et al. [26] developed a 2,2,6,6-tetramethylpiperidine-1-oxyl-4-amino-4-carboxylic acid spin-labeled amyloid-β (Aβ) peptide to observe its isoform-dependent interaction with the ApoE protein. Oligomer binding involves the C-terminal domain of ApoE, with ApoE3 reporting a much greater response through this conformational marker. ApoE3 displays a higher affinity and capacity for the toxic Aβ oligomer. ApoE polymorphism and AD risk can largely be attributed to the reduced ability of ApoE4 to function as a clearance vehicle for the toxic form of Aβ. *MAPT* and *APOE* are involved in the pathogenic mechanisms of AD, and both the *MAPT H1/H1* genotype and the *APOE ε4* allele lead to a more rapid progression to dementia among MCI subjects, probably mediating an increased rate of amyloid-β and tau brain deposition [27].

The distribution of *APOE* genotypes in the Iberian peninsula is as follows: *APOE-2/2* 0.32%; *APOE-2/3* 7.3%; *APOE-2/4* 1.27%; *APOE-3/3* 71.11%; *APOE-3/4* 18.41%; and *APOE-4/4* 1.59% [18] (Figure 27.2). These frequencies are very similar in Europe and in other Western societies. There is a clear accumulation of *APOE-4* carriers among patients with AD (*APOE-3/4* 30.30%, *APOE-4/4* 6.06%) and vascular dementia (*APOE-3/4* 35.85%, *APOE-4/4* 6.57%) as compared to controls (Figure 27.2). Different *APOE* genotypes confer specific phenotypic profiles to AD patients [15,16,22]. Some of these profiles may add risk or benefit when patients are treated with conventional drugs, and in many instances the clinical phenotype demands the administration of additional drugs that increase the complexity of therapeutic protocols.

From studies designed to define *APOE*-related AD phenotypes [7,12,23–25,28,53–59], several conclusions can be drawn, which are shown in Box 27.2. These 20 major phenotypic features clearly illustrate the biological disadvantage of *APOE-4* homozygotes and the potential consequences that these patients may experience when they receive pharmacological treatment for AD and/or concomitant pathologies [7,12,23–25,28,53–59].

27.7.6 Pharmacogenomics of Antidementia Drugs

The following list describes the pharmacogenomics of the most common antidementia drugs (Table 27.1).

Donepezil: is a centrally active, reversible acetylcholinesterase inhibitor that increases the acetylcholine available for synaptic transmission in the CNS. The therapeutic response of donepezil is influenced by pathogenic gene variants (*APOE*, *CHAT*), as well as mechanistic gene polymorphic variants (*CHAT*, *ACHE*, and *BCHE*). It is a major substrate of CYP2D6, CYP3A4, ACHE, and UGTs; it inhibits ACHE and BCHE; and it is transported by ABCB1 [113].

Galantamine: is a reversible and competitive acetylcholinesterase inhibitor leading to increased concentration of acetylcholine at cholinergic synapses. It also modulates nicotinic acetylcholine receptors and may increase glutamate and serotonin levels. *APOE*, *APP*, *ACHE*, *BCHE*, *CHRNA4*, *CHRNA7*, and *CHRNB2* variants may potentially influence galantamine efficacy and safety. Galantamine is a major substrate of CYP2D6, CYP3A4, and UGT1A1, and an inhibitor of ACHE and BCHE [113].

BOX 27.2 Key Conclusions Regarding APOE-Related AD Phenotypes

1. The age at onset is 5–10 years earlier in approximately 80% of AD cases harboring the *APOE-4/4* genotype.
2. The serum levels of ApoE are lowest in *APOE-4/4*, intermediate in *APOE-3/3* and *APOE-3/4*, and highest in *APOE-2/3* and *APOE-2/4*.
3. Serum cholesterol levels are higher in *APOE-4/4* than in other genotypes.
4. HDL-cholesterol levels tend to be lower in *APOE-3* homozygotes than in *APOE-4* allele carriers.
5. LDL-cholesterol levels are systematically higher in *APOE-4/4* than in any other genotype.
6. Triglyceride levels are significantly lower in *APOE-4/4*.
7. Nitric oxide levels are slightly lower in *APOE-4/4*.
8. Serum and cerebrospinal fluid Aβ levels tend to differ between *APOE-4/4* and the other most frequent genotypes (*APOE-3/3*, *APOE-3/4*).
9. Blood histamine levels are dramatically reduced in *APOE-4/4* as compared to the other genotypes.
10. Brain atrophy is markedly increased in *APOE-4/4*>*APOE-3/4*>*APOE-3/3*.
11. Brain mapping activity shows a significant increase in slow wave activity in *APOE-4/4* from the early stages of the disease.
12. Brain hemodynamics, as reflected by reduced brain blood flow velocity and increased pulsatility and resistance indices, is significantly worse in *APOE-4/4* (and in *APOE-4* carriers in general, as compared with *APOE-3* carriers); brain hypoperfusion and neocortical oxygenation is also more deficient in *APOE-4* carriers.
13. Lymphocyte apoptosis is markedly enhanced in *APOE-4* carriers.
14. Cognitive deterioration is faster in *APOE-4/4* patients than in carriers of any other *APOE* genotype.
15. In approximately 3–8% of AD cases, some dementia-related metabolic dysfunctions accumulate more in *APOE-4* carriers than in *APOE-3* carriers.
16. Some behavioral disturbances, alterations in circadian rhythm patterns, and mood disorders are slightly more frequent in *APOE-4* carriers.
17. Aortic and systemic atherosclerosis is more frequent in *APOE-4* carriers.
18. Liver metabolism and transaminase activity differ in *APOE-4/4* with respect to other genotypes.
19. Hypertension and other cardiovascular risk factors accumulate in *APOE-4* carriers.
20. *APOE-4/4* carriers are the poorest responders to conventional drugs.

Rivastigmine: is a cholinesterase inhibitor that increases acetylcholine in the CNS through reversible inhibition of its hydrolysis by cholinesterase. *APOE*, *APP*, *CHAT*, *ACHE*, *BCHE*, *CHRNA4*, *CHRNB2*, and *MAPT* variants may affect its pharmacokinetics and pharmacodynamics [113].

Tacrine: is the first FDA-approved antidementia drug. Its use was stopped due to hepatotoxicity. Tacrine is a cholinesterase inhibitor that elevates acetylcholine in the cerebral cortex by slowing degradation of acetylcholine. *ACHE*, *BCHE*, *CHRNA4*, *CHRNB2*, *APOE*, *MTHFR*, *CES1*, *LEPR*, *GSTM1*, and *GSTT1* variants may affect its therapeutic and toxic effects. Tacrine is a major substrate of CYP1A2 and CYP3A4, a minor substrate of CYP2D6, and is transported via SCN1A. It is an inhibitor of ACHE, BCHE, and CYP1A2 [113].

Memantine: is an N-Methyl-D-Aspartate (NMDA) receptor antagonist that binds preferentially to NMDA receptor-operated cation channels. It may act by blocking the actions of glutamate, mediated in part by NMDA receptors, and it is also an antagonist of GRIN2A, GRIN2B, GRIN3A, HTR3A, and CHRFAM7A. Several pathogenic (*APOE*, *PSEN1*, *MAPT*) and mechanistic gene variants (*GRIN2A*, *GRIN2B*, *GRIN3A*, *HTR3A*, *CHRFAM7A*) may influence its therapeutic effects. Memantine is a strong inhibitor of CYP2B6 and CYP2D6, and a weak inhibitor of CYP1A2, CYP2A6, CYP2C9, CYP2C19, CYP2E1, and CYP3A4 [113].

27.7.7 Multifactorial Therapy

Some studies using a multifactorial approach also have shown that diverse pharmacogenomic factors may influence efficacy and safety. In one of these studies [15,58], patients with dementia received the following for three months: a multifactorial therapy integrated by CDP-choline (500 mg/day, p.o.), Nicergoline (5 mg/day, p.o.), Sardilipin (E-SAR-94010) (LipoEsar®)(250 mg, t.i.d.), and Animon Complex® (2 capsules/day)—a nutraceutical compound integrated by a purified extract of *Chenopodium quinoa* (250 mg), ferrous sulphate (38.1 mg equivalent to 14 mg of iron), folic acid (200 µg), and vitamin B$_{12}$ (1 µg) per capsule (RGS: 26.06671/C).

Patients with chronic deficiencies of iron (<35 µg/mL), folic acid (<2.5 ng/mL), or vitamin B$_{12}$ (<150 pg/mL) received an additional supplement of iron (80 mg/day), folic acid (5 mg/day), and B complex vitamins (B$_1$, 15 mg/day; B$_2$, 15 mg/day; B$_6$, 10 mg/day; B$_{12}$, 10 µg/day; nicotinamide, 50 mg/day), respectively, to maintain stable levels of serum iron (50–150 µg/mL), folic acid (5–20 ng/mL) and vitamin B$_{12}$ levels (500–1000 pg/mL) in order to avoid the negative influence of all these metabolic factors on cognition. Patients with hypertension (>150/85 mmHg) received Enalapril (20 mg/day).

The frequency of *APOE* genotypes was *APOE-2/3*, 7.97%; *APOE-2/4*, 1.18%; *APOE-3*, 58.95%; *APOE-3/4*, 27.32%; and *APOE-4/4*, 4.58%. Cognitive function (as assessed by MMSE); 20.51 ± 6.51 vs. 21.45 ± 6.95, $p<0.0000000001$; ADAS-Cog, 22.94 ± 13.87 vs. 21.23 ± 12.84, $p<0.0001$; ADAS-Non-Cog, 5.26 ± 4.18 vs. 4.15 ± 3.63, $p<0.0000000001$; ADAS-Total, 27.12 ± 16.93 vs. 24.28 ± 15.06, $p<0.00009$) improved after treatment. Mood (HAM-A, 11.35 ± 5.44 vs. 9.79 ± 4.33, $p<0.0000000001$; HAM-D, 10.14 ± 5.23 vs. 8.59 ± 4.30, $p<0.0000000001$) also improved. Glucose levels did not change.

Total cholesterol levels (224.78 ± 45.53 vs. 203.64 ± 39.69 mg/dL, $p<0.0000000001$), HDL-cholesterol levels (54.11 ± 14.54 vs. 52.54 ± 14.86 mg/dL, $p<0.0001$), and LDL-cholesterol levels (148.15 ± 39.13 vs. 128.89 ± 34.83 mg/dL, $p<0.0000000001$) were significantly reduced. Folate (7.07 ± 3.61 vs. 18.14 ± 4.23 ng/mL, $p<0.000000001$) and vitamin B$_{12}$ levels (459.65 ± 205.80 vs. 689.78 ± 338.82 pg/mL, $p<0.000000001$) also increased, and both TSH and T$_4$ levels remained unchanged after treatment. The response rate in terms of cognitive improvement was as follows: 59.74% responders (RRs), 24.44% nonresponders (NRs), and 15.82% stable responders (SRs) (no change in MMSE score after three months of treatment). The response rate in cholesterol levels was very similar: 57.78% RRs, 28.50% NRs, and 13.72% SRs [15].

27.7.7.1 *APOE*-Related Cognitive Function Changes

In this study, the basal MMSE score differed in *APOE-2/3* carriers with respect to *APOE-2/4* ($p<0.02$), *APOE-3/4* ($p<0.004$), and *APOE-4/4* ($p<0.0009$), in *APOE-3/3* vs. *APOE-3/4* ($p<0.0005$), and in *APOE-3/3* vs. *APOE-4/4* ($p<0.002$). The best responders were *APOE-3/3* ($p<0.0000000001$) >*APOE-3/4* ($p<0.00001$) >*APOE-4/4* carriers ($p<0.05$). Patients harboring the *APOE-2/3* and *APOE-2/4* genotypes did not show any significant improvement. The response rate by genotype was the following: *APOE-2/3*: 44.26% RRs, 36.07% NRs, 19.67% SRs; *APOE-2/4*: 55.56% RRs, 44.44% NRs, 0.0% SRs; *APOE-3/3*: 63.42% RRs, 21.06% NRs, 15.52% SRs; *APOE-3/4*: 56.94% RRs, 27.75% NRs, 15.31% SRs; and *APOE-4/4*: 51.43% RRs, 28.57% NRs, 20.00% SRs [15] (Figures 27.6 and 27.7).

27.7.7.2 *APOE*-Related Changes in Blood Pressure Values

Systolic blood pressure (SBP) was significantly reduced in patients with the *APOE-3/3* ($p<0.00007$) and *APOE-3/4* genotypes ($p<0.01$), and diastolic blood pressure exhibited a similar pattern (*APOE-3/3*, $p<0.005$; *APOE-3/4*,

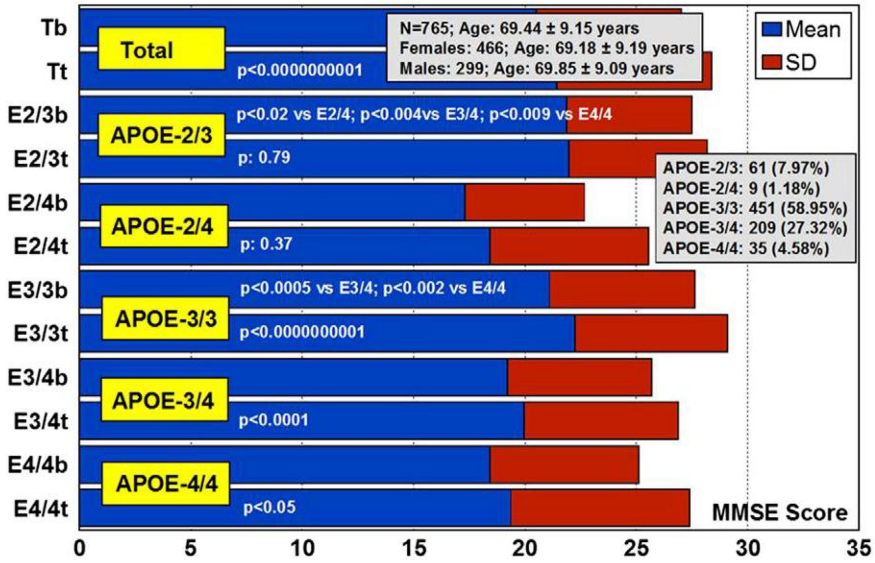


FIGURE 27.6 ***APOE*-related cognitive performance in response to multifactorial therapy in patients with dementia.** Tb—basal MMSE score prior to treatment; Tt—MMSE score after 3 months treatment in total sample. E2/3b—basal MMSE score in *APOE-2/3* carriers; E2/3t—MMSE score after treatment in *APOE-2/3* carriers; E2/4b—basal MMSE score in *APOE-2/4* carriers; E2/4t—MMSE score after treatment in *APOE-2/4* carriers; E3/3b—basal MMSE score in *APOE-3/3* carriers; E3/3t—MMSE score after treatment in *APOE-3/3* carriers; E3/4b basal MMSE score in *APOE-3/4* carriers; E3/4t—MMSE score after treatment in *APOE-3/4* carriers; E4/4b—basal MMSE score in *APOE-4/4* carriers; E4/4—MMSE score after treatment in *APOE-4/4* carriers. *Source: Adapted from Cacabelos et al.* [15].

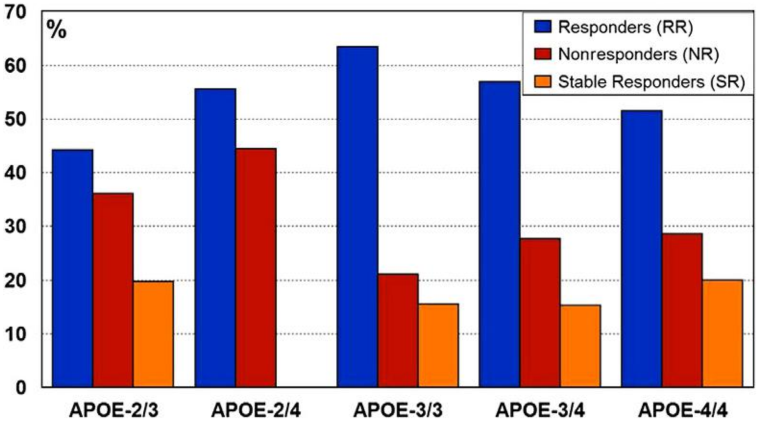


FIGURE 27.7 ***APOE*-related cognitive response rate in patients with dementia treated with multifactorial therapy.**

p < 0.01), with no changes in either SBP or DBP in *APOE-2/3*, *APOE-2/4*, and *APOE-4/4* carriers [15].

27.7.7.3 APOE-Related Blood Lipid Response to Sardilipin

Basal cholesterol levels were significantly different in patients with the *APOE-2/3* genotype vs. *APOE-3/3* (p < 0.007), vs. *APOE-3/4* (p < 0.001), vs. *APOE-4/4* (p < 0.00002); *APOE-2/4* vs. *APOE-4/4* (p < 0.01); *APOE-3/3* vs. *APOE-4/4* (p < 0.005); and *APOE-3/4* vs. *APOE-4/4* (p < 0.01).

The highest cholesterol levels were seen in *APOE-4/4*> *APOE-3/4*> *APOE-3/3*. All patients showed a clear reduction in cholesterol levels after treatment with Sardilipin. This was particularly significant in *APOE-3/3* (p < 0.0000000001) > *APOE-3/4* (p < 0.00000008) > *APOE-4/4* (p < 0.002) > *APOE-2/3* (p < 0.02) > *APOE-2/4* carriers (p: 0.26). The response rate by genotype was as follows: *APOE-2/3*: 63.93% RRs, 29.51% NRs, 6.56% SRs; *APOE-2/4*: 44.44% RRs, 22.22% NRs, 33.34% SRs; *APOE-3/3*: 54.32% RRs, 28.16% NRs, 17.52% SRs; *APOE-3/4*: 53.59% RRs, 31.58% NRs, 14.83% SRs; *APOE-4/4*: 65.71% RRs, 20.00% NRs, 14.29% SRs [15].

HDL-cholesterol levels significantly decreased in *APOE-3/3* ($p<0.001$) > *APOE-3/4* ($p<0.05$), with no significant changes in patients with other genotypes. In contrast, LDL-cholesterol levels showed changes identical to those observed in total cholesterol, with similar differences among genotypes at baseline and almost identical decreased levels after treatment (*APOE-3/3*, $p>0.0000000001$ >*APOE-3/4*, $p<0.00001$ >*APOE-2/3*, $p<0.0004$ >*APOE-4/4*, $p<0.001$ > *APOE-2/4*, p:0.31) [15].

Sardilipin (E-SAR-94010, LipoEsar®, LipoSea®) is a natural product extracted from the marine species *Sardina pilchardus* by means of nondenaturing biotechnological procedures. The main chemical compounds of LipoEsar® are lipoproteins (60–80%), whose micelle structure probably mimics that of physiological lipoproteins involved in lipid metabolism. In preclinical studies, Sardilipinhas been shown to be effective in:

1. Reducing blood cholesterol (CHO), triglyceride (TG), uric acid (UA), and glucose (Glu) levels, as well as liver alanine aminotransferase (ALT) and aspartate aminotransferase (AST) activity.
2. Enhancing immunological function by regulating both lymphocyte and microglia activity.
3. Inducing antioxidant effects mediated by superoxide dismutase activity.
4. Improving cognitive function [15].

According to these results, it appears that the therapeutic response of patients with dyslipidemia to Sardilipin is *APOE*-related. The best responders were patients with *APOE-3/3* >*APOE-3/4* >*APOE-4/4*. Patients with the other *APOE* genotypes (*2/2*, *2/3*, *2/4*) did not show any hypolipidemic response to this novel compound. In patients with dementia, the effects of Sardilipin were very similar to those observed in patients with chronic dyslipidemia, suggesting that the lipid-lowering properties of Sardilipin are *APOE*-dependent [15] (Figure 27.8).

27.8 FUTURE PERSPECTIVE

To make AD a global health priority in the coming years, conceptual and procedural changes are needed on several grounds, such as (1) political, administrative, economic, legal, ethical, industrial, regulatory and educational issues; (2) novel biomarkers (genomics, proteomics, molecular neuroimaging) as diagnostic aids; (3) innovative therapeutics; (4) pharmacogenomics in clinical practice to optimize therapeutics; and (5) selective preventive plans for the population at risk.

There is disharmony concerning the public and governmental interest in dementia and its social, medical, and economic implications. The diagnosis and management of dementia is dissimilar in Europe, North America, Latin America, Asia, Africa, and Oceania. The economic/cultural status of each country (developed versus developing), the

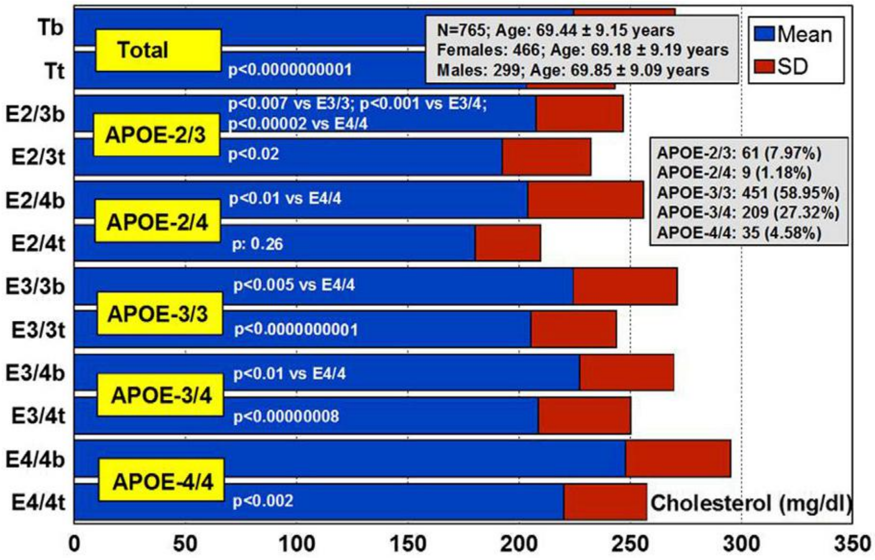


FIGURE 27.8 ***APOE*-related total cholesterol levels in response to multifactorial therapy in patients with dementia.** Tb—basal cholesterol levels prior to treatment; Tt—cholesterol levels after 3 months treatment in total sample. E2/3b—basal cholesterol levels in *APOE-2/3* carriers; E2/3t—total cholesterol levels after treatment in *APOE-2/3* carriers; E2/4b—basal cholesterol levels in *APOE-2/4* carriers; E2/4t—total cholesterol levels after treatment in *APOE-2/4* carriers; E3/3b—basal cholesterol levels in *APOE-3/3* carriers; E3/3t—total cholesterol levels after treatment in *APOE-3/3* carriers; E3/4b—basal cholesterol levels in *APOE-3/4* carriers; E3/4t—total cholesterol levels after treatment in *APOE-3/4* carriers; E4/4b—basal cholesterol levels in *APOE-4/4* carriers; E4/4—total cholesterol levels after treatment in *APOE-4/4* carriers. *Source: Adapted from Cacabelos et al.* [15].

particular epidemiology of aging and dementia in each latitude, national standards of education, health priorities (infectious diseases versus degenerative diseases), and the quality and efficiency of medical services are conditioning factors for investing (or not investing) national resources in dementia as a health priority.

Educational programs, international guidelines, and consensus protocols for the management of dementia are necessary for global harmonization, for professionals from different countries to speak the same conceptual language, and to improve cost-effectiveness ratios [125–128]. There are many legal issues (e.g., informed consent, lawsuit, testament, tutorship) and ethical issues (e.g., clinical trials, use of genetic information, institutionalization) that deserve more attention in order to humanize the end of life in the very frail conditions under which demented patients survive.

The updating of regulatory issues is also a matter of deep concern. Regulatory aspects of drug development are not universal, with notable peculiarities in the European Union (EMA), the United States (FDA), and Japan (Koseisho). Because the costs of dementia cannot be fully assumed by more than 60% of the European population, European authorities must take into account this circumstance when health reform is implemented in the coming years [8,18].

Genomics, transcriptomics, proteomics, and metabolomics will revolutionize medicine in the next decades. Genetic testing is gaining acceptance among physicians and patients in different countries [128–131], although Americans, Europeans, and Japanese differ notably in their knowledge, beliefs, and attitudes regarding genetic testing for AD [128,131,132]. The validation of protocols for genomic screening will contribute to the implementation of structural genomics, functional genomics, and proteomics as diagnostic aids and therapeutic targets [133].

An accurate diagnosis of AD demands the use of reliable biomarkers in routine protocols at a reasonable price [68]. Levels of specific secreted cellular signaling proteins in cerebrospinal fluid or plasma correlate with pathological changes in the AD brain; therefore, proteomic analysis of these levels can be used to discover said biomarkers [134]. It is likely that the best biomarkers result from a combination of genomic, transcriptomic, and proteomic analyses of body fluids. The measurement of these biomarkers correlates with brain imaging markers and cognitive performance [73–75].

New initiatives for the prevention of dementia (global versus selective prevention) will also emerge [135], together with new insights into the role of nutrition and nutrigenomics in brain function and neurodegeneration [59,136]. In terms of prevention, it must be taken into consideration that neuronal death and Aβ accumulation starts many years before the onset of the disease, and that preventive strategies should be selective to protect the population at risk. For this purpose, accurate biomarkers are essential, and surrogate markers are needed to facilitate primary prevention.

Without doubt, the highest priority for the coming decade will be an intense search for novel therapeutic options in the form of both symptomatic treatments and preventive strategies. Past failures must be studied by researchers and the pharmaceutical industry in order to avoid unnecessary expenses in redundant trials that lead nowhere. Combination treatments require further evaluation and more sophisticated strategies than dual combinations [137,138]. The administration of psychotropic drugs to demented patients should be reduced and predicted with pharmacogenetic markers to minimize side effects, cerebrovascular risk, and cognitive deterioration.

Priority areas for pharmacogenetic research are the prediction of serious adverse drug reactions (ADRs) and the determination of efficacy variation [139]. Both are necessary in CNS disorders and dementia to cope with efficacy and safety issues associated with current psychotropics and antidementia drugs, as well as new CNS drugs. With regard to the future of pharmacogenomics as a practical discipline, several issues should be addressed:

- The education of physicians in medical genomics and pharmacogenomics is fundamental (less than 2% of clinicians are familiar with genomic science)
- Genomic screening of gene clusters involved in pharmacogenomic outcomes must become a clinical routine (without genetic testing, there is no pharmacogenetics)
- Each patient must be a carrier of a pharmacogenetic card [140] indicating what kind of drugs he/she can take and which medications he/she should avoid
- Regulatory agencies should request pharmacogenetic data from the pharmaceutical industry when applying for drug approval
- Pharmacogenetic data must be incorporated into patient information leaflets and the pharmaceutical *vade mecum*
- New guidelines for daily praxis, such as those given in the *World Guide for Drug Use and Pharmacogenomics* [113], will promote understanding of the relationship between drugs and genes to make drug prescription truly personalized

27.9 CONCLUSION

AD is a major health problem that comes with a high cost to society. As a clinical entity, AD is a polygenic/complex disorder in which many different gene clusters may be involved. Most genes screened to date belong to different proteomic and metabolomic pathways that potentially affect AD pathogenesis, represented by accumulation of Aβ deposits in senile plaques, intracellular NFTs with hyperphosphorylated tau, and neuronal loss.

The presence of the *APOE-4* allele of the apolipoprotein E gene seems to be a major risk factor for both degenerative and vascular dementia, and *APOE* variants are directly involved in AD pathogenesis at multiple levels. Specific biomarkers (structural and functional genomic markers, proteomic markers in body fluids, neuroimaging markers) are needed for an accurate AD diagnosis. Current pharmacological treatment of AD with cholinesterase inhibitors (donepezil, rivastigmine, galantamine) and memantine is not cost-effective; moreover, the overuse of psychotropic drugs in patients with dementia contributes to deteriorating cognitive and psychomotor functions.

Old treatments addressed memory impairment. New treatments are oriented to halting disease progression by interfering with Aβ accumulation, NFT formation, oxidative stress, neuroinflammation, and cerebrovascular damage. Over the past few years diverse candidate drugs have been investigated in AD models, but not one has reached the market. Since only 25–30% of the population is an extensive metabolizer for drugs metabolized via CYP2D6, CYP2C9, and CYP2C19 enzymes, it seems reasonable to use pharmacogenomic procedures as a way to optimize AD therapeutics, thus reducing ADRs and unnecessary costs. The therapeutic response to conventional drugs in patients with AD is genotype-specific, with *CYP2D6*-PMs, *CYP2D6*-UMs, and *APOE-4/4* carriers shown to be the worst responders. *APOE* and *CYP2D6* may cooperate, as pleiotropic genes, in the metabolism of drugs and hepatic function.

If we know the pharmacogenomic profiles of patients who require treatment with antidementia drugs and/or psychotropic drugs currently in use, we may be able to achieve the following benefits:

- Identifying candidate patients with the ideal genomic profile to receive a particular drug
- Adapting the dose in more than 90% of cases according to the condition of EM, IM, PM, or UM, which will limit the occurrence of direct side effects in 30–50% of cases
- Reducing drug interactions by 30–50% (avoiding the administration of inhibitors or inducers able to modify the normal enzymatic activity on a particular substrate)
- Enhancing efficacy
- Eliminating unnecessary costs (>30% of pharmaceutical direct costs) deriving from the consequences of inappropriate drug selection and overmedication to mitigate ADRs [18]

APPENDIX A

Selected Genes Potentially Associated with Alzheimer's Disease

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
1p13.1	52.32	NGF	Nerve growth factor (beta polypeptide)	162030	Hereditary sensory and autonomic neuropathy type V, allergic rhinitis
1p13.3	5.95	GSTM1	Glutathione S-transferase mu 1	138350	Cancer
1p13.3	7.11	GSTM3	Glutathione S-transferase mu 3 (brain)	138390	Cancer
1p13.3	20.38	CSF1	Colony-stimulating factor 1 (macrophage)	120420	
1p13-p12	20.94	HMGCS2	3-hydroxy-3-methylglutaryl-CoA synthase 2 (mitochondrial)	600234	HMG-CoA synthase-2 deficiency
1p21	232.03	COL11A1	Collagen, type XI, alpha 1	120280	Fibrochondrogenesis, Marshall syndrome, Stickler syndrome type II, lumbar disc herniation
1p21.3-p13.1	88.38	SORT1	Sortilin 1	602458	
1p22.2	18.49	GBP2	Guanylate binding protein 2, interferon-inducible	600412	
1p31.3	44.38	TM2D1	TM2-domain containing 1	610080	

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
1p32	134.20	ERI3	ERI1 exoribonuclease family member 3	609917	
1p32.3	37.62	DHCR24	24-dehydrocholesterol reductase	606418	Desmosterolosis
1p32.3	204.31	ZFYVE9	Zinc finger, FYVE-domain containing 9	603755	
1p34	85.79	LRP8	Low-density lipoprotein receptor-related protein 8, apolipoprotein e receptor	602600	Myocardial infarction, major depressive disorder
1p34.3	34.93	LCK	Lymphocyte-specific protein tyrosine kinase	153390	SCID due to LCK deficiency
1p36.1	128.30	ECE1	Endothelin-converting enzyme 1	600423	Hirschsprung disease, cardiac defects, autonomic dysfunction, essential hypertension
1p36.13-q31.3	3.81	APH1A	APH1A gamma secretase subunit provided	607629	
1p36.1-p34	115.01	HSPG2	Heparan sulfate proteoglycan 2	142461	Dyssegmental dysplasia Silverman-Handmaker type, Schwartz-Jampel syndrome type 1, tardive dyskinesia
1p36.22	12.87	TARDBP	TAR DNA binding protein	605078	Amyotrophic lateral sclerosis 10 with or without FTD, frontotemporal lobar degeneration TARDBP-related
1p36.3	83.64	TP73	Tumor protein p73	601990	Neuroblastoma
1p36.3	20.37	MTHFR	Methylenetetrahydrofolate reductase (NAD(P)H)	607093	Homocystinuria due to MTHFR deficiency, neural tube defects, schizophrenia, thromboembolism, occlusive vascular disease, colon cancer, acute leukemia
1p36-p35	14.28	HTR6	5-Hydroxytryptamine (serotonin) receptor 6, G protein-coupled	601109	
1q21	35.76	CTSS	Cathepsin S	116845	
1q21	N/A	AD13	Alzheimer disease 13	611152	
1q21	3.64	S100A1	S100 calcium binding protein A1	176940	Cardiomyopathies
1q21.3	11.55	FAM63A	Family with sequence similarity 63, member A	N/A	
1q21.3	12.10	CHRNB2	Cholinergic receptor, nicotinic, beta 2 (neuronal)	118507	Nocturnal frontal lobe epilepsy 3
1q21-q22	66.10	NTRK1	Neurotrophic tyrosine kinase, receptor, type 1	191315	Insensitivity to pain with anhidrosis, medullary thyroid carcinoma, self-mutilating behavior, mental retardation

(Continued)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
1q21-q23	1.05	APCS	Amyloid P component, serum	104770	Secondary amyloidosis
1q22	57.51	LMNA	Lamin A/C	150330	Emery-Dreifuss muscular dystrophy 2, Emery-Dreifuss muscular dystrophy 3, familial partial lipodystrophy 2, muscular dystrophy, limb girdle muscular dystrophy type 1B, dilated cardiomyopathy 1A, Charcot-Marie-Tooth disease type 2B1, Hutchinson-Gilford progeria syndrome, heart-hand syndrome of Slovenian type, Malouf syndrome, mandibuloacral dysplasia, lethal restrictive dermopathy
1q22	11.92	FDPS	Farnesyl diphosphate synthase	134629	
1q22-q23	15.68	NCSTN	Nicastrin	605254	Acne inversa 1
1q22-q23	6.72	USF1	Upstream transcription factor 1	191523	Hyperlipidemia
1q23	3.95	FCER1G	Fc fragment of IgE, high-affinity I, receptor forgamma polypeptide	147139	
1q23.2	2.30	CRP	C-reactive protein, pentraxin-related provided	123260	
1q24.2	206.52	POU2F1	POU class 2 homeobox 1	164175	
1q25	64.97	SOAT1	Sterol O-acyltransferase 1	102642	
1q25	N/A	AD14	Alzheimer disease 14	611154	
1q25.2-q25.3	8.62	PTGS2	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	600262	
1q31-q32	4.89	IL10	Interleukin 10	124092	Rheumatoid arthritis
1q31-q42	25.53	PSEN2	Presenilin 2 (Alzheimer disease 4)	600759	Dilated cardiomyopathy 1V
1q32	95.63	CFH	Complement factor H	134370	Hemolytic-uremic syndrome, chronic hypocomplementemic nephropathy, basal laminar drusen, complement factor H deficiency, macular degeneration 4
1q32	145.64	CR1	Complement component (3b/4b) receptor 1 (Knops blood group)	120620	CR1 deficiency, systemic lupus erythematosus
1q32-q41	48.77	HSD11B1	Hydroxysteroid (11-beta) dehydrogenase 1	600713	Cortisone reductase deficiency 2, obesity, insulin resistance

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
1q41-q42	47.41	PARP1	Poly (ADP-ribose) polymerase 1	173870	Xeroderma pigmentosum, Fanconi anemia, diabetes type I
1q42.2	12.07	AGT	Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	106150	Renal tubular dysgenesis, non-familial structural atrial fibrillation, inflammatory bowel disease, essential hypertension, preeclampsia
1q43	108.70	MTR	5-methyltetrahydrofolate-homocysteine methyltransferase	156570	Methylcobalamin deficiency type cblG, neural tube defects
2p12-p11.1	1140	CTNNA2	Catenin (cadherin-associated protein), alpha 2	114025	
2p16.3	78.41	RTN4	Reticulon 4	604475	
2p21	68.97	LHCGR	Luteinizing hormone/choriogonadotropin receptor	152790	Leydig cell adenoma with precocious puberty, Leydig cell hypoplasia with hypergonadotropic hypogonadism, Leydig cell hypoplasia with pseudohermaphroditism, female luteinizing hormone resistance, male precocious puberty
2p22-p21	51.91	EIF2AK2	Eukaryotic translation initiation factor 2-alpha kinase 2	176871	
2p25	66.53	ADAM17	ADAM metallopeptidase domain 17	603639	Neonatal inflammatory skin and bowel disease
2q14	11.48	IL1A	Interleukin 1, alpha	147760	Rheumatoid arthritis
2q14	7.02	IL1B	Interleukin 1, beta	147720	
2q14	59.31	BIN1	Bridging integrator 1	601248	Centronuclear myopathy
2q14.2	16.12	IL1RN	Interleukin 1 receptor antagonist	147679	Interleukin 1 receptor antagonist deficiency, microvascular complications of diabetes 4
2q21.1	88.75	KCNIP3	Kv channel interacting protein 3, calsenilin	604662	
2q21.2	1900	LRP1B	Low-density lipoprotein receptor-related protein 1B	608766	
2q24-q31	235.50	LRP2	Low-density lipoprotein receptor-related protein 2	600073	Donnai-Barrow syndrome, facio-oculoacousticorenal syndrome
2q34	75.67	CREB1	cAMP responsive element binding protein 1	123810	Angiomatoid fibrous histiocytoma
3p21.31	7.18	CCR2	Chemokine (C-C motif) receptor 2	601267	
3p21.31	6.07	CCR5	Chemokine (C-C motif) receptor 5 (gene/pseudogene)	601373	Insulin-dependent diabetes mellitus 22

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
3p25	146.51	PPARG	Peroxisome proliferator-activated receptor gamma	601487	Carotid intimal medial thickness 1, insulin resistance, lipodystrophy 3, obesity, diabetes type 2, cancer
3p26.2	16.73	OGG1	8-oxoguanine DNA glycosylase	601982	Renal cell carcinoma
3q13.3	272.47	GSK3B	Glycogen synthase kinase 3 beta	605004	Parkinson disease
3q21.3	88.66	RAB7A	RAB7A, member RAS oncogene family	602298	Charcot-Marie-Tooth disease type 2B
3q22.1	32.87	TF	Transferrin	190000	Atransferrinemia
3q22-q24	N/A	AD15	Alzheimer disease 15	611155	
3q25.2	104.08	MME	Membrane metallo-endopeptidase	120520	Membranous glomerulonephritis, neutral endopeptidase deficiency
3q26.1-q26.2	64.56	BCHE	Butyrylcholinesterase	177400	
3q26.2-qter	15.50	APOD	Apolipoprotein D	107740	
3q27	8.26	AHSG	Alpha-2-HS-glycoprotein	138680	
3q28	1.51	SST	Somatostatin	182450	
4p13	404.59	APBB2	Amyloid beta (A4) precursor protein binding, family B, member 2	602710	
4p14	11.55	UCHL1	Ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	191342	Parkinson disease 5
4p14-p13	78.93	RFC1	Replication factor C (activator 1) 1, 145kDa	102579	
4p16.1	550.19	SORCS2	Sortilin-related VPS10 domain-containing receptor 2	606284	
4p16.3	28.90	LRPAP1	Low-density lipoprotein receptor-related protein-associated protein 1	104225	
4q13.3	17.16	ALB	Albumin	103600	Analbuminemia, dysalbuminemic hyperthyroxinemia, dysalbuminemic hyperzincemia
4q13-q21	3.21	IL8	Interleukin 8	146930	Bronchiolitis
4q21	114.20	SNCA	Synuclein, alpha (non-A4 component of amyloid precursor)	163890	Lewy body dementia, Parkinson disease 1, Parkinson disease 4
4q25	491.92	COL25A1	Collagen, type XXV, alpha 1	610004	
4q25	14.85	CASP6	Caspase 6, apoptosis-related cysteine peptidase	601532	
4q27	29.00	ANXA5	Annexin A5	131230	Recurrent pregnancy loss

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
4q32	21.80	TLR2	Toll-like receptor 2	603028	Colorectal cancer
4q32.1	9.11	LRAT	Lecithin retinol acyltransferase (phosphatidylcholine-retinol O-acyltransferase)	604863	Leber congenital amaurosis 14, retinal dystrophy, retinitis pigmentosa
5p15.3	52.64	SLC6A3	Solute carrier family 6 (neurotransmitter transporter, dopamine), member 3	126455	Epilepsy, parkinsonism-dystonia, attention-deficit hyperactivity disorder, Parkinson disease
5q13.1	86.07	PIK3R1	Phosphoinositide-3-kinase, regulatory subunit 1 (alpha)	171833	Agammaglobulinemia 7, insulin resistance
5q13.3-q14	24.93	HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase	142910	
5q14.1	209.33	ARSB	Arylsulfatase B	611542	Mucopolysaccharidosis type VI (Maroteaux-Lamy)
5q15	112.65	CAST	Calpastatin	114090	
5q21	294.01	EFNA5	Ephrin-A5	601535	
5q31	105.89	FGF1	Fibroblast growth factor 1 (acidic)	131220	
5q31	6.34	APBB3	Amyloid beta (A4) precursor protein-binding, family B, member 3	602711	
5q31.1	1.97	CD14	CD14 molecule	158120	
5q31-q32	2.04	ADRB2	Adrenoceptor beta 2, surface	109690	Nocturnal asthma, obesity, diabetes type 2
5q32	491.97	PPP2R2B	Protein phosphatase 2, regulatory subunit B, beta	604325	Spinocerebellar ataxia 12
5q34	180.24	WWC1	WW and C2 domain–containing 1	610533	
5q35.3	17.08	DBN1	Drebrin 1	126660	
6p12	16.28	VEGFA	Vascular endothelial growth factor A	192240	Microvascular complications of diabetes 1
6p12	149.48	CD2AP	CD2-associated protein	604241	Focal segmental glomerulosclerosis 3
6p21	38.96	GLP1R	Glucagon-like peptide 1 receptor	138032	
6p21.1	4.68	TREM2	Triggering receptor expressed on myeloid cells 2	605086	Nasu-Hakola disease
6p21.3	7.96	HFE	Hemochromatosis	613609	Hemochromatosis, microvascular complications of diabetes 7, porphyria cutanea tarda, porphyria variegata
6p21.3	4.31	UBD	Ubiquitin D	606050	

(Continued)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
6p21.3	3.42	HLA-A	Major histocompatibility complex, class I, A	142800	
6p21.3	12.26	DDX39B	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39B	142560	Rheumatoid arthritis
6p21.3	2.77	TNF	Tumor necrosis factor	191160	Asthma, vascular dementia, migraine without aura, insulin resistance, cancer
6p21.3	2.43	HSPA1A	Heat shock 70kDa protein 1A	140550	
6p21.3	16.91	PPP1R10	Protein phosphatase 1, regulatory subunit 10	603771	
6p21.3	3.36	AGER	Advanced glycosylation end product-specific receptor	600214	Diabetes
6p21.3	16.94	TAP2	Transporter 2, ATP-binding cassette, subfamily B (MDR/TAP)	170261	Bare lymphocyte syndrome type I due to TAP2 deficiency, Wegener-like granulomatosis, ankylosing spondylitis, insulin-dependent diabetes mellitus, celiac disease
6p21.3	47.89	C2	Complement component 2	613927	C2 deficiency
6p21.3	20.62	C4B	Complement component 4B (Chido blood group)	120820	C4B deficiency, systemic lupus erythematosus
6p21.3	20.62	C4A	Complement component 4A (Rodgers blood group)	120810	C4A deficiency, systemic lupus erythematosus, type I diabetes mellitus
6p21.3	13.05	MICB	MHC class I polypeptide-related sequence B	602436	
6p21.3	7.11	RXRB	Retinoid X receptor, beta	180246	
6p21.33	11.72	MICA	MHC class I polypeptide-related sequence A	600169	
6p22.1	21.01	PGBD1	PiggyBac transposable element derived 1	N/A	
6p23	462.38	ATXN1	Ataxin 1	601556	Spinocerebellar ataxia 1
6p25.3-p24.3	176.61	F13A1	Coagulation factor XIII, A1 polypeptide	134570	Factor XIIIA deficiency
6p25-p24	199.05	NEDD9	Neural precursor cell expressed, developmentally downregulated 9	602265	Cancer metastasis
6q21	213.12	FYN	FYN oncogene related to SRC, FGR, YES	137025	
6q21	49.75	SNX3	Sorting nexin 3	605930	
6q25.1	412.78	ESR1	Estrogen receptor 1	133430	Breast cancer, atherosclerosis, migraine, myocardial infarction, endometrial cancer, osteoporosis

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
6q25.3	14.21	SOD2	Superoxide dismutase 2, mitochondrial	147460	Microvascular complications of diabetes 6, cardiomyopathy, premature aging, sporadic motor neuron disease, cancer
6q27	18.54	TBP	TATA box binding protein	600075	Spinocerebellar ataxia 17, Parkinson disease
7p15.1	7.68	NPY	Neuropeptide Y	162640	Elevated cholesterol levels, higher alcohol consumption, metabolic diseases, cardiovascular diseases
7p21	4.86	IL6	Interleukin 6 (interferon, beta 2)	147620	Crohn disease-associated growth failure, diabetes, intracranial hemorrhage in brain cerebrovascular malformations, Kaposi sarcoma, rheumatoid arthritis
7p22	8.92	NUDT1	Nudix- (nucleoside diphosphate-linked moiety X)-type motif 1	600312	
7q11.2	77.09	CD36	CD36 molecule (thrombospondin receptor)	173510	Platelet glycoprotein IV deficiency, macrothrombocytopenia, coronary heart disease
7q21	1440	MAGI2	Membrane-associated guanylate kinase, WW and PDZ domain-containing 2	606382	
7q21.12	209.46	ABCB1	ATP-binding cassette, subfamily B (MDR/TAP), member 1	171050	Inflammatory bowel disease 13
7q21.3	26.22	PON1	Paraoxonase 1	168820	Coronary artery disease, coronary artery spasm 2, microvascular complications of diabetes 5
7q21.3	36.50	PON3	Paraoxonase 3	602720	
7q21.3	30.21	PON2	Paraoxonase 2	602447	Coronary artery disease, diabetes
7q22	517.73	RELN	Reelin	600514	Lissencephaly 2 (Norman-Roberts type)
7q22.1	12.18	SERPINE1	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	173360	Plasminogen activator inhibitor-1 deficiency, thrombophilia
7q31.1	42.20	PPP1R3A	Protein phosphatase 1, regulatory subunit 3A	600917	Insulin resistance
7q31.1	36.40	CAV1	Caveolin 1, caveolae protein, 22kDa	601047	Lipodystrophy type 3
7q31-q32	30.06	DLD	Dihydrolipoamide dehydrogenase	238331	Dihydrolipoamide dehydrogenase deficiency, maple syrup urine disease

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
7q34	17.78	EPHA1	EPH receptor A1	179610	Cancer
7q36	23.54	NOS3	Nitric oxide synthase 3 (endothelial cell)	163729	Coronary artery spasm 1, hypertension, ischemic stroke, placental abruption
7q36	4.15	CDK5	Cyclin-dependent kinase 5	123831	
7q36	59.28	PAXIP1	PAX-interacting (with transcription, activation domain) protein 1	608254	
7q36	N/A	AD10	Alzheimer disease-10	609636	
8p11.2	8.38	STAR	Steroidogenic acute regulatory protein	600617	Lipoid adrenal hyperplasia
8p11.22	108.28	ADAM9	ADAM metallopeptidase domain 9	602713	Cone-rod dystrophy 9
8p12	3.67	ADRB3	Adrenoceptor beta 3	109691	Obesity
8p12	32.96	PLAT	Plasminogen activator, tissue	173370	Hyperfibrinolysis, thrombophilia
8p12	29.86	EIF4EBP1	Eukaryotic translation initiation factor 4E binding protein 1	602223	
8p12-q22	N/A	AD12	Alzheimer disease 12	611073	
8p21-p12	17.90	CLU	Clusterin	185430	Neoplasms
8p22	9.97	NAT2	N-acetyltransferase 2 (arylamine N-acetyltransferase)	612182	Cancer
8p22	28.19	LPL	Lipoprotein lipase	609708	Hyperlipidemia, lipoprotein lipase deficiency
8p22	25.61	CTSB	Cathepsin B	116810	Esophageal adenocarcinoma, neoplasms
8q13	2.24	CRH	Corticotropin releasing hormone	122560	
8q22	87.63	DPYS	Dihydropyrimidinase	613326	Dihydropyrimidinuria
8q24.1	81.79	ENPP2	Ectonucleotide pyrophosphatase/phosphodiesterase 2	601060	
9p13.3	3.05	SIGMAR1	Sigma nonopioid intracellular receptor 1	601978	Amyotrophic lateral sclerosis 16
9p13.3	16.68	VCP	Valosin-containing protein	601023	Amyotrophic lateral sclerosis 14 with or without frontotemporal dementia, inclusion body myopathy with early-onset Paget disease and frontotemporal dementia
9p21	26.74	CDKN2A	Cyclin-dependent kinase inhibitor 2A	600160	Melanoma and neural system tumor syndrome, orolaryngeal cancer, pancreatic cancer, cutaneous malignant melanoma 2

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
9p21.3	126.31	CDKN2B-AS1	CDKN2B antisense RNA 1	613149	Cardiovascular diseases, cancer, intracranial aneurysm, type-2 diabetes, periodontitis, endometriosis, frailty in the elderly, glaucoma
9p24	32.69	VLDLR	Very low-density lipoprotein receptor	192977	Cerebellar hypoplasia and mental retardation with or without quadrupedal locomotion 1
9p24.1	42.20	IL33	Interleukin 33	608678	
9q13-q21.1	244.83	APBA1	Amyloid beta (A4) precursor protein-binding, family A, member 1	602414	
9q21.2	294.71	PRUNE2	Prune homolog 2 (*Drosophila*)	610691	
9q21.2-q21.3	48.29	UBQLN1	Ubiquilin 1	605046	Parkinson disease
9q21.33	210.79	DAPK1	Death-associated protein kinase 1	600831	
9q21.33	74.06	GOLM1	Golgi membrane protein 1	606804	
9q22.1	355.04	NTRK2	Neurotrophic tyrosine kinase, receptor, type 2	600456	Obesity, mood disorders
9q22.1	N/A	AD11	Alzheimer disease 11	609790	
9q31.1	169.23	GRIN3A	Glutamate receptor, ionotropic, N-methyl-D-aspartate 3A	606650	
9q31.1	147.25	ABCA1	ATP-binding cassette, subfamily A (ABC1), member 1	600046	HDL deficiency type 2, Tangier disease, coronary artery disease
9q33.1	13.32	TLR4	Toll-like receptor 4	603030	Colorectal cancer, macular degeneration
9q33.3	6.54	HSPA5	Heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	138120	
9q34	22.98	DBH	Dopamine beta-hydroxylase (dopamine beta-monooxygenase)	609312	Dopamine beta-hydroxylase deficiency
9q34	40.10	TRAF2	TNF receptor-associated factor 2	601895	
9q34	21.69	ABCA2	ATP-binding cassette, subfamily A (ABC1), member 2	600047	
9q34.3	114.12	RXRA	Retinoid X receptor, alpha	180245	
10	51.74	ENTPD7	Ectonucleoside triphosphate diphosphohydrolase 7	N/A	

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
10p12	401.08	CACNB2	Calcium channel, voltage-dependent, beta 2 subunit	600003	Brugada syndrome 4
10p12.31	161.34	C10orf112	Chromosome 10 open reading frame 112	N/A	
10p13	38.20	OPTN	Optineurin	602432	Amyotrophic lateral sclerosis 12, glaucoma
10p13	N/A	AD7	Alzheimer disease 7	606187	
10p14-p13	27.42	PTPLA	Protein tyrosine phosphatase-like (proline instead of catalytic arginine), member A	610467	
10p15.2	35.12	PITRM1	Pitrilysin metallopeptidase 1	N/A	
10q	N/A	AD6	Alzheimer disease 6	605526	
10q11.2	71.94	ALOX5	Arachidonate 5-lipoxygenase	152390	Atherosclerosis, cancer
10q11.2	56.01	CHAT	Choline O-acetyltransferase	118490	Myasthenic syndrome associated with episodic apnea
10q11.2	3.38	DKK1	Dickkopf WNT signaling pathway inhibitor 1	605189	
10q21	14.09	TFAM	Transcription factor A, mitochondrial	600438	
10q21	707.23	ANK3	Ankyrin 3, node of Ranvier (ankyrin G)	600465	
10q21	134.12	TET1	Tet methylcytosine dioxygenase 1	607790	
10q21	65.24	SGPL1	Sphingosine-1-phosphate lyase 1	603729	
10q21.1	16.52	CDK1	Cyclin-dependent kinase 1	116940	
10q21.3	175.08	LRRTM3	Leucine-rich repeat transmembrane neuronal 3	610869	
10q21.3	33.72	SIRT1	Sirtuin 1	604479	
10q22.2	1780	CTNNA3	Catenin (cadherin-associated protein), alpha 3	607667	
10q22.2	6.40	PLAU	Plasminogen activator, urokinase	191840	Quebec platelet disorder
10q22.3	768.22	KCNMA1	Potassium large conductance calcium-activated channel, subfamily M, alpha member 1	600150	Generalized epilepsy and paroxysmal dyskinesia
10q23	1.38	CH25H	Cholesterol 25-hydroxylase	604551	
10q23.2-q23.3	38.34	LIPA	Lipase A, lysosomal acid, cholesterol esterase	613497	Cholesteryl ester storage disease, Wolman disease

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
10q23.3	105.34	PTEN	Phosphatase and tensin homolog	601728	Bannayan-Riley-Ruvalcaba syndrome, Cowden syndrome 1, endometrial carcinoma, Lhermitte-Duclos syndrome, macrocephaly/autism syndrome, malignant melanoma, PTEN hamartoma tumor syndrome, squamous cell carcinoma, thyroid carcinoma, VATER association with macrocephaly and ventriculomegaly, glioma, meningioma, prostate cancer
10q23.32	104.42	HECTD2	HECT domain-containing E3 ubiquitin protein ligase 2	N/A	
10q23.33	5.73	HHEX	Hematopoietically expressed homeobox	604420	
10q23-q25	122.41	IDE	Insulin-degrading enzyme	146680	Type 2 diabetes mellitus
10q23-q25	624.13	SORCS3	Sortilin-related VPS10 domain-containing receptor 3	606285	
10q23-q25	591.05	SORCS1	Sortilin-related VPS10 domain-containing receptor 1	606283	
10q24	23.92	COX15	Cytochrome c oxidase assembly homolog 15 (yeast)	603646	Cardioencephalomyopathy due to cytochrome c oxidase deficiency 2, Leigh syndrome due to cytochrome c oxidase deficiency
10q24	69.20	ABCC2	ATP-binding cassette, subfamily C (CFTR/MRP), member 2	601107	Dubin-Johnson syndrome
10q24.1	25.25	FAS	Fas cell surface death receptor	134637	Autoimmune lymphoproliferative syndrome type IA, squamous cell carcinoma, autoimmune lymphoproliferative syndrome
10q24.2	134.34	DNMBP	Dynamin binding protein	611282	
10q24.3	50.88	ALDH18A1	Aldehyde dehydrogenase 18 family, member A1	138250	Cutis laxa type IIIA, hyperammonemia, hypoornithinemia, hypocitrullinemia, hypoarginemia, hypoprolinemia, cataracts, connective tissue diseases
10q24.3	7.00	CYP17A1	Cytochrome P450, family 17, subfamily A, polypeptide 1	609300	17,20-lyase deficiency, 17-alpha-hydroxylase/17,20-lyase deficiency, pseudohermaphroditism, adrenal hyperplasia
10q24.31	17.82	SCD	Stearoyl-CoA desaturase (delta-9-desaturase)	604031	

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
10q24.33	261.38	SH3PXD2A	SH3 and PX domains 2A	N/A	
10q24.33	5.50	CALHM1	Calcium homeostasis modulator 1	612234	
10q25.1	13.27	GSTO1	Glutathione S-transferase omega 1	605482	
10q25.1	30.55	GSTO2	Glutathione S-transferase omega 2	612314	
10q25.3	2.86	ADRB1	Adrenoceptor beta 1	109630	Heart failure
10q26.3	128.60	EBF3	Early B-cell factor 3	607407	Glioblastoma multiforme, gastric carcinoma
10q26.3	53.38	HTRA1	HtrA serine peptidase 1	602194	CARASIL syndrome, macular degeneration
10q26.3	374.23	ADAM12	ADAM metallopeptidase domain 12	602714	
11p11.2	20.97	MAPK8IP1	Mitogen-activated protein kinase 8 interacting protein 1	604641	Noninsulin-dependent diabetes mellitus
11p13	67.17	BDNF	Brain-derived neurotrophic factor	113505	Central hypoventilation syndrome, anorexia nervosa, bulimia nervosa, memory impairment, obsessive-compulsive disorder
11p15	24.29	APBB1	Amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)	602709	
11p15.1	3.72	SAA1	Serum amyloid A1	104750	Atherosclerosis, rheumatoid arthritis, Crohn's disease, neoplasms
11p15.5	3.40	DRD4	Dopamine receptor D4	126452	Autonomic nervous system dysfunction, novelty-seeking personality, attention-deficit hyperactivity disorder
11p15.5	11.24	CTSD	Cathepsin D	116840	Breast cancer, neuronal ceroid lipofuscinosis type 10
11p15.5	1.43	INS	Insulin	176730	Insulin-dependent diabetes mellitus type 2, permanent neonatal diabetes mellitus, diabetes mellitus type 1, familial hyperproinsulinemia with or without diabetes
11p15.5	1.59	HBG2	Hemoglobin, gamma G	142250	Transient neonatal cyanosis
11q12.1	13.06	MS4A6A	Membrane-spanning 4-domains, subfamily A, member 6A	606548	
11q12.2	41.84	MS4A4E	Membrane-spanning 4-domains, subfamily A, member 4E	608401	
11q13	3.06	GSTP1	Glutathione S-transferase pi 1	134660	Cancer

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
11q13	14.31	INPPL1	Inositol polyphosphate phosphatase-like 1	600829	Opsismodysplasia, breast cancer
11q13.3	6.66	GAL	Galanin/GMAP prepropeptide	137035	
11q14	112.71	PICALM	Phosphatidylinositol-binding clathrin assembly protein	603025	Acute myeloid leukemia, T-cell acute lymphoblastic leukemia, malignant lymphomas
11q14.1	202.53	GAB2	GRB2-associated binding protein 2	606203	
11q22.2-q22.3	25.73	CASP4	Caspase 4, apoptosis-related cysteine peptidase	602664	
11q22.2-q22.3	20.87	IL18	Interleukin 18 (interferon-gamma-inducing factor)	600953	
11q22.3	8.33	MMP1	Matrix metallopeptidase 1 (interstitial collagenase)	120353	Epidermolysis bullosa dystrophica, arthritis, chronic obstructive pulmonary disease
11q22.3	7.82	MMP3	Matrix metallopeptidase 3 (stromelysin 1, progelatinase)	185250	Coronary heart disease, arthritis
11q23	3.05	APOA5	Apolipoprotein A-V	606368	Hyperchylomicronemia, hypertriglyceridemia, hyperlipoproteinemia type V
11q23	2.59	APOA4	Apolipoprotein A-IV	107690	
11q23.2-q23.3	30.57	BACE1	Beta-site APP-cleaving enzyme 1	604252	
11q23.2-q24.2	181.56	SORL1	Sortilin-related receptor, L(DLR class) A repeats containing	602005	
11q23.3	3.16	APOC3	Apolipoprotein C-III	107720	Hyperalphalipoproteinemia 2, hypertriglyceridemia
11q23-q24	1.87	APOA1	Apolipoprotein A-I	107680	Amyloidosis, combined ApoA-I and apoC-III deficiency, corneal clouding, hypoalphalipoproteinemia, Tangier disease, systemic non-neuropathic amyloidosis
11q24	74.99	APLP2	Amyloid beta (A4) precursor-like protein 2	104776	
11q25	12.32	ACAD8	Acyl-CoA dehydrogenase family, member 8	604773	Isobutyryl-CoA dehydrogenase deficiency
12p11.23-q13.12	N/A	AD5	Alzheimer disease 5	602096	
12p12	418.61	GRIN2B	Glutamate receptor, ionotropic, N-methyl D-aspartate 2B	138252	Mental retardation
12p12.1	7.11	IAPP	Islet amyloid polypeptide	147940	Diabetes type 2
12p13	4.55	PKP2P1	Plakophilin 2 pseudogene 1	602861	Arrhythmogenic right ventricular dysplasia 9

(Continued)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
12p13	3.95	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	138400	
12p13	7.18	GNB3	Guanine nucleotide binding protein (G protein), beta polypeptide 3	139130	Essential hypertension, obesity
12p13.2	150.85	LRP6	Low-density lipoprotein receptor-related protein 6	603507	Coronary artery disease
12p13.2-p12.3	13.89	OLR1	Oxidized low density lipoprotein (lectin-like) receptor 1	602601	Myocardial infarction, atherosclerosis
12p13.3	37.84	NCAPD2	Non-SMC condensin I complex, subunit D2	N/A	
12p13.31	10.31	TAPBPL	TAP binding protein-like	607081	
12p13.31	48.26	A2M	Alpha-2-macroglobulin	103950	Alpha-2-macroglobulin deficiency
12q12	144	LRRK2	Leucine-rich repeat kinase 2	609007	Parkinson disease 8
12q13	79.39	TFCP2	Transcription factor CP2	189889	
12q13	118.56	ATF7	Activating transcription factor 7	606371	
12q13.11	63.49	VDR	Vitamin D (1,25-dihydroxyvitamin D3) receptor	601769	Involutional osteoporosis
12q13.11	29.04	KANSL2	KAT8 regulatory NSL complex subunit 2	N/A	
12q13.11	28.54	CCNT1	Cyclin T1	143055	Neoplasms
12q13.3	84.86	LRP1	Low-density lipoprotein receptor-related protein 1	107770	
12q14	45.33	CAND1	Cullin-associated and neddylation-dissociated 1	607727	
12q23-q24.1	5.68	PLA2G1B	Phospholipase A2, group IB (pancreas)	172410	
12q24.11	61.42	KIAA1033	KIAA1033	N/A	
12q24.2	43.10	ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)	100650	Esophageal cancer alcohol-related
12q24.2-q24.31	153.66	NOS1	Nitric oxide synthase 1	163731	Stroke
13q22.1	18.54	KLF5	Kruppel-like factor 5 (intestinal)	602903	
13q34	25.17	DAOA	D-amino acid oxidase activator	607408	Schizophrenia, bipolar affective disorder
14q13.1	8.63	PNP	Purine nucleoside phosphorylase	164050	Immunodeficiency due to purine nucleoside phosphorylase deficiency
14q21	114.25	SOS2	Son of sevenless homolog 2 (*Drosophila*)	601247	

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
14q22.1	241.61	FRMD6	FERM domain-containing 6	614555	
14q23.2	111.52	ESR2	Estrogen receptor 2 (ER beta)	601663	
14q24	71.97	MTHFD1	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase	172460	Abruptio placentae, spina bifida folate-sensitive
14q24.3	87.26	PSEN1	Presenilin 1	104311	Familial acne inversa 3, dilated cardiomyopathy 1U, frontotemporal dementia, Pick disease
14q24.3	13.44	NPC2	Niemann-Pick disease, type C2	601015	Niemann-Pick disease type C2
14q24.3	21.86	DLST	Dihydrolipoamide S-succinyltransferase (E2 component of 2-oxo-glutarate complex)	126063	
14q24.3	3.46	FOS	FBJ murine osteosarcoma viral oncogene homolog	164810	
14q24.3	5.82	NGB	Neuroglobin	605304	
14q31	62.31	SEL1L	Sel-1 suppressor of lin-12-like (*C. elegans*)	602329	
14q32	76.35	MOK	MOK protein kinase	605762	
14q32.1	13.95	SERPINA1	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	107400	Emphysema due to AAT deficiency, emphysema-cirrhosis due to AAT deficiency, hemorrhagic diathesis due to antithrombin Pittsburgh, chronic obstructive pulmonary disease
14q32.1	11.68	SERPINA3	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	107280	Alpha-1-antichymotrypsin deficiency, occlusive cerebrovascular disease
14q32.1	42.88	CYP46A1	Cytochrome P450, family 46, subfamily A, polypeptide 1	604087	
14q32.3	72.36	KLC1	Kinesin light chain 1	600025	
14q32.32	26.40	AKT1	V-akt murine thymoma viral oncogene homolog 1	164730	Breast cancer, colorectal cancer, Cowden syndrome 6, ovarian cancer, proteus syndrome, schizophrenia
15q11-q12	196.68	APBA2	Amyloid beta (A4) precursor protein-binding, family A, member 2	602712	

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
15q13.1	32.42	CHRFAM7A	CHRNA7 (cholinergic receptor, nicotinic, alpha 7, exons 5–10) and FAM7A (family with sequence similarity 7A, exons A–E) fusion	609756	
15q14	139.70	CHRNA7	Cholinergic receptor, nicotinic, alpha 7 (neuronal)	118511	Schizophrenia, myoclonic epilepsy
15q21.1	130.54	CYP19A1	Cytochrome P450, family 19, subfamily A, polypeptide 1	107910	Aromatase deficiency, aromatase excess syndrome
15q21.1	2.09	EID1	EP300 interacting inhibitor of differentiation 1	605894	
15q21-q23	136.90	LIPC	Lipase, hepatic	151670	Hepatic lipase deficiency, noninsulin-dependent diabetes mellitus
15q22	153.67	ADAM10	ADAM metallopeptidase domain 10	602192	
15q22.2	31.58	APH1B	APH1B gamma secretase subunit	607630	
15q22.31	48.35	SNX1	Sorting nexin 1	601272	
15q22.33	129.34	SMAD3	SMAD family member 3	603109	
15q24	28.24	CHRNA3	Cholinergic receptor, nicotinic, alpha 3 (neuronal)	118503	Lung cancer
15q24.1	21.11	CSK	C-src tyrosine kinase	124095	
15q25.1	63.28	IREB2	Iron-responsive element binding protein 2	147582	
15q26	150.50	MEF2A	Myocyte enhancer factor 2A	600660	Coronary artery disease 1 with myocardial infarction
16p13.3	17.87	UBE2I	Ubiquitin-conjugating enzyme E2I	601661	
16p13.3	14.60	MEFV	Mediterranean fever	608107	Familial Mediterranean fever
16q12	29.56	VPS35	Vacuolar protein sorting 35 homolog (*S. cerevisiae*)	601501	Parkinson disease 17
16q21	21.92	CETP	Cholesteryl ester transfer protein, plasma	118470	Hyperalphalipoproteinemia
16q22	28.10	NAE1	NEDD8-activating enzyme E1 subunit 1	603385	
16q22.1	17.23	NQO1	NAD(P)H dehydrogenase, quinone 1	125860	Tardive dyskinesia, cancer
17p11.2	25.66	SREBF1	Sterol regulatory element-binding transcription factor 1	184756	

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
17p12	139.28	COX10	Cytochrome c oxidase assembly homolog 10 (yeast)	602125	Charcot-Marie-Tooth type 1A, hereditary neuropathy with liability to pressure palsies
17p13	72.14	MYH13	Myosin, heavy-chain 13, skeletal muscle	603487	
17p13.1	8.80	TNK1	Tyrosine kinase, nonreceptor, 1	608076	
17p13.1	19.15	TP53	Tumor protein p53	191170	Adrenal cortical carcinoma, breast cancer, choroid plexus papilloma, colorectal cancer, hepatocellular carcinoma, Li-Fraumeni syndrome, nasopharyngeal carcinoma, osteosarcoma, pancreatic cancer, basal-cell carcinoma 7, glioma
17p13.1	31.63	MYH8	Myosin, heavy-chain 8, skeletal muscle, perinatal	160741	Carney complex variant, trismus-pseudocamptodactyly syndrome
17q	2.22	PNMT	Phenylethanolamine N-methyltransferase	171190	Essential hypertension
17q11.2	4.17	CDK5R1	Cyclin-dependent kinase 5, regulatory subunit 1 (p35)	603460	
17q11.2	43.97	BLMH	Bleomycin hydrolase	602403	
17q11.2	39.58	SLC6A4	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	182138	Sudden infant death syndrome, aggressive behavior, depression, obsessive-compulsive disorder
17q11.2	31.68	THRA	Thyroid hormone receptor, alpha	190120	Hypothyroidism nongoitrous 6
17q11.2	0,086	MIR144	MicroRNA 144	612070	
17q11.2-q12	43.76	NOS2	Nitric oxide synthase 2, inducible	163730	Hypertension
17q11.2-q12	1.93	CCL2	Chemokine (C-C motif) ligand 2	158105	Psoriasis, rheumatoid arthritis, atherosclerosis, spina bifida
17q12	1.91	CCL3	Chemokine (C-C motif) ligand 3	182283	Human immunodeficiency virus type 1
17q21.1	133.95	MAPT	Microtubule-associated protein tau	157140	Pick disease, frontotemporal dementia, cortico-basal degeneration, progressive supranuclear palsy, Parkinson disease, tauopathy and respiratory failure
17q21.1	0,445	STH	Saitohin	607067	

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
17q21.32	7.98	GRN	Granulin	138945	Frontotemporal lobar degeneration with ubiquitin-positive inclusions, primary progressive aphasia, neuronal ceroid lipofuscinosis 11
17q21-q22	N/A	GPSC	Gliosis, familial progressive subcortical	N/A	
17q23.1	11.08	MPO	Myeloperoxidase	606989	Myeloperoxidase deficiency
17q23.2	83.06	APPBP2	Amyloid beta precursor protein (cytoplasmic tail) binding protein 2	605324	
17q23.3	21.32	ACE	Angiotensin I converting enzyme (peptidyl-dipeptidase A) 1	106180	Renal tubular dysgenesis, benign serum increase of angiotensin I-converting enzyme, myocardial infarction, stroke, severe acute respiratory syndrome, microvascular complications of diabetes 3
17q24.3	158.72	BPTF	Bromodomain PHD finger transcription factor	601819	
17q24-q25	87.63	GRB2	Growth factor receptor-bound protein 2	108355	
18p11.2	33.49	MC2R	Melanocortin 2 receptor (adrenocorticotropic hormone)	607397	Glucocorticoid deficiency due to ACTH unresponsiveness
18q12.1	33.62	DSC1	Desmocollin 1	125643	
18q12.1	7.26	TTR	Transthyretin	176300	Amyloidotic polyneuropathy, euthyroid hyperthyroxinaemia, amyloidotic vitreous opacities, cardiomyopathy, oculoleptomeningeal amyloidosis, meningocerebrovascular amyloidosis, carpal tunnel syndrome
19p13	14.48	PIN1	Peptidylprolyl *cis/trans* isomerase, NIMA-interacting 1	601052	Cancer
19p13.2	113.86	DNM2	Dynamin 2	602378	Axonal Charcot-Marie-Tooth disease type 2M, Charcot-Marie-Tooth disease type B, centronuclear myopathy
19p13.2	44.47	LDLR	Low-density lipoprotein receptor	606945	Familial hypercholesterolemia
19p13.2	N/A	AD9	Alzheimer disease 9	608907	
19p13.2-p13.1	41.35	NOTCH3	Notch 3	600276	Cerebral arteriopathy with subcortical infarcts and leukoencephalopathy
19p13.3-13.2	42.82	C3	Complement component 3	120700	C3 deficiency, atypical hemolytic uremic syndrome, macular degeneration age-related 9

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
19p13.3	27.06	GNA11	Guanine nucleotide binding protein (G protein), alpha 11 (Gq class)	139313	
19p13.3	25.47	ABCA7	ATP-binding cassette, subfamily A (ABC1), member 7	605414	
19p13.3	10.90	APBA3	Amyloid beta (A4) precursor protein-binding, family A, member 3	604262	
19p13.3	9.29	GRIN3B	Glutamate receptor, ionotropic, N-methyl-D-aspartate 3B	606651	
19p13.3-p13.2	181.75	INSR	Insulin receptor	147670	Diabetes mellitus insulin-resistant with acanthosis nigricans, hyperinsulinemic hypoglycemia 5, leprechaunism, Rabson-Mendenhall syndrome
19p13.3-p13.2	15.78	ICAM1	Intercellular adhesion molecule 1	147840	
19q13	12.47	TOMM40	Translocase of outer mitochondrial membrane 40 homolog (yeast)	608061	
19q13.1	23.02	TGFB1	Transforming growth factor, beta 1	190180	Camurati-Engelmann disease
19q13.1	11.30	APLP1	Amyloid beta (A4) precursor-like protein 1	104775	
19q13.12	11.91	GAPDHS	Glyceraldehyde-3-phosphate dehydrogenase, spermatogenic	609169	
19q13.12	1.41	PSENEN	Presenilin enhancer 2 homolog (*C. elegans*)	607632	Acne inversa 2
19q13.2	43.09	PVRL2	Poliovirus receptor-related 2 (herpesvirus entry mediator B)	600798	Multiple sclerosis
19q13.2	3.61	APOE	Apolipoprotein E	107741	Hyperlipoproteinemia type III, lipoprotein glomerulopathy, sea-blue histiocyte disease, macular degeneration, myocardial infarction
19q13.2	4.69	APOC1	Apolipoprotein C-I	107710	
19q13.2	3.26	APOC4	Apolipoprotein C-IV	600745	Coronary artery disease
19q13.2	3.58	APOC2	Apolipoprotein C-II	608083	Hyperlipoproteinemia type Ib
19q13.2	12.36	BCAM	Basal cell adhesion molecule (Lutheran blood group)	612773	
19q13.3	38.76	CLPTM1	Cleft lip and palate-associated transmembrane protein 1	612585	

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
19q13.3	14.94	CD33	CD33 molecule	159590	
19q13.3	6.61	NR1H2	Nuclear receptor subfamily 1, group H, member 2	600380	
19q13.3	54.03	MARK4	MAP/microtubule affinity-regulating kinase 4	606495	
19q13.32	21.59	EXOC3L2	Exocyst complex component 3-like 2	N/A	
19q13.32	3.06	BLOC1S3	Biogenesis of lysosomal organelles complex1, subunit 3	609762	Hermansky-Pudlak syndrome 8
19q13.33	47.86	CARD8	Caspase recruitment domain family, member 8	609051	Rheumatoid arthritis
19q13.43	9.76	GALP	Galanin-like peptide	611178	Neuroblastic tumor
20p	N/A	AD8	Alzheimer disease 8	607116	
20p11.21	4.28	CST3	Cystatin C	604312	Cerebral amyloid angiopathy, macular degeneration 11
20p13	15.44	PRNP	Prion protein	176640	Creutzfeldt-Jakob disease, fatal familial insomnia, Gerstmann-Straussler disease, Huntington disease-like 1, kuru, prion disease
20pter-p12	6.55	PRND	Prion protein 2 (dublet)	604263	
20q13.2-q13.3	18.09	CHRNA4	Cholinergic receptor, nicotinic, alpha 4 (neuronal)	118504	Nocturnal frontal lobe epilepsy type 1
20q13.31	5.38	PCK1	Phosphoenolpyruvate carboxykinase 1 (soluble)	614168	Cytosolic phosphoenolpyruvate carboxykinase deficiency
21q11	98.17	SAMSN1	SAM domain, SH3 domain, and nuclear localization signals 1	607978	
21q21.1	134.54	TMPRSS15	Transmembrane protease, serine 15	606635	Enterokinase deficiency
21q21.1	541.88	NCAM2	Neural cell adhesion molecule 2	602040	
21q21.3	290.59	APP	Amyloid beta (A4) precursor protein	104760	Cerebral amyloid angiopathy
21q22.1	291.96	KCNJ6	Potassium inwardly rectifying channel, subfamily J, member 6	600877	
21q22.11	102.95	EVA1C	Eva-1 homolog C (*C. elegans*)	N/A	
21q22.11	3.79	DNAJC28	DnaJ (Hsp40) homolog, subfamily C, member 28	N/A	
21q22.13	147.82	DYRK1A	Dual-specificity tyrosine-(Y)-phosphorylation-regulated kinase 1A	600855	Down syndrome, mental retardation 7

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
21q22.2	129.73	DOPEY2	Dopey family member 2	604803	
21q22.3	261.50	RUNX1	Runt-related transcription factor 1	151385	Acute myeloid leukemia, platelet disorder with associated myeloid malignancy
21q22.3	108.80	BACE2	Beta-site APP-cleaving enzyme 2	605668	Down syndrome
21q22.3	97.56	ABCG1	ATP-binding cassette, subfamily G (WHITE), member 1	603076	
21q22.3	23.17	CBS	Cystathionine-beta-synthase	613381	Cystathionine beta-synthase deficiency, homocystinuria, hyperhomocysteinemic thrombosis
21q22.3	50.19	MCM3AP	Minichromosome maintenance complex component 3-associated protein	603294	
21q22.3	6.50	S100B	S100 calcium binding protein B	176990	Down syndrome, epilepsy, amyotrophic lateral sclerosis, melanoma, type I diabetes
22q11.21	28.24	COMT	Catechol-O-methyltransferase	116790	Panic disorder, schizophrenia
22q11.21	26.88	RTN4R	Reticulon 4 receptor	605566	Schizophrenia
22q11.23	0.845	MIF	Macrophage migration inhibitory factor (glycosylation-inhibiting factor)	153620	Rheumatoid arthritis
22q11.23	8.15	GSTT1	Glutathione S-transferase theta 1	600436	Carcinoma
22q13.1	13.15	HMOX1	Heme oxygenase (decycling) 1	141250	Heme oxygenase-1 deficiency, chronic obstructive pulmonary disease
22q13.2	21.30	SEPT3	Septin 3	608314	
22q13.31	93.16	PPARA	Peroxisome proliferator-activated receptor alpha	170998	Hyperapobetalipoproteinemia
Xp11.2	3.12	HSD17B10	Hydroxysteroid (17-beta) dehydrogenase 10	300256	17-beta-hydroxysteroid dehydrogenase X deficiency, 2-methyl-3-hydroxybutyryl-CoA dehydrogenase deficiency, mental retardation
Xp11.23	115.86	MAOB	Monoamine oxidase B	309860	
Xp11.3	91.92	MAOA	Monoamine oxidase A	309850	Brunner syndrome
Xp11.3-p11.23	4.50	TIMP1	TIMP metallopeptidase inhibitor 1	305370	
Xp21.1	68.97	OTC	Ornithine carbamoyltransferase	300461	Ornithine transcarbamylase deficiency, Duchenne muscular dystrophy

(*Continued*)

Appendix A Continued

Locus	Size (Kb)	Symbol	Title/Gene	OMIM	Other Related Diseases
Xq12	186.59	AR	Androgen receptor	313700	Spinal and bulbar muscular atrophy of Kennedy, prostate cancer, complete androgen insensitivity, hypospadia type 1
Xq21.3	843.97	PCDH11X	Protocadherin 11 X-linked	300246	
Xq21.3	N/A	AD16	Alzheimer disease 16	300756	

APPENDIX B

Long-Form Names for Genes Listed in Table 27.2

ADH1A: Alcohol dehydrogenase 1A (class I), alpha polypeptide
AADAC: Arylacetamide deacetylase
AANAT: Aralkylamine N-acetyltransferase
ACSL1: Acyl-CoA synthetase long-chain family member 1
ACSL3: Acyl-CoA synthetase long-chain family member 3
ACSL4: Acyl-CoA synthetase long-chain family member 4
ACSM1: Acyl-CoA synthetase medium-chain family member 1
ACSM2B: Acyl-CoA synthetase medium-chain family member 2B
ACSM3: Acyl-CoA synthetase medium-chain family, member 3
ADH1B: Alcohol dehydrogenase 1B (class I), beta polypeptide
ADH1C: Alcohol dehydrogenase 1C (class I), gamma polypeptide
ADH4: Alcohol dehydrogenase 4 (class II), pi polypeptide
ADH5: Alcohol dehydrogenase 5 (class III), chi polypeptide
ADH6: Alcohol dehydrogenase 6 (class V)
ADH7: Alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide
ADHFE1: Alcohol dehydrogenase, iron containing, 1
AGXT: Alanine-glyoxylate aminotransferase
AKR1A1: Aldo-keto reductase family 1, member A1 (aldehyde reductase)
AKR1B1: Aldo-keto reductase family 1, member B1 (aldose reductase)
AKR1C1: Aldo-keto reductase family 1, member C1
AKR1D1: Aldo-keto reductase family 1, member D1
ALDH1A1: Aldehyde dehydrogenase 1 family, member A1
ALDH1A2: Aldehyde dehydrogenase family 1, subfamily A2
ALDH1A3: Aldehyde dehydrogenase family 1, subfamily A3
ALDH1B1: Aldehyde dehydrogenase 1 family, member B1
ALDH2: Aldehyde dehydrogenase 2 family (mitochondrial)
ALDH3A1: Aldehyde dehydrogenase 3 family, member A1
ALDH3A2: Aldehyde dehydrogenase 3 family, member A2
ALDH3B1: Aldehyde dehydrogenase 3 family, member B1
ALDH3B2: Aldehyde dehydrogenase 3 family, member B2
ALDH4A1: Aldehyde dehydrogenase 4 family, member A1
ALDH5A1: Aldehyde dehydrogenase 5 family, member A1
ALDH6A1: Aldehyde dehydrogenase 6 family, member A1
ALDH7A1: Aldehyde dehydrogenase 7 family, member A1
ALDH8A1: Aldehyde dehydrogenase 8 family, member A1
ALDH9A1: Aldehyde dehydrogenase 9 family, member A1
AOX1: Aldehyde oxidase 1
AS3MT: Arsenic (+3 oxidation state) methyltransferase
ASMT: Acetylserotonin O-methyltransferase
BAAT: Bile acid CoA: amino acid N-acyltransferase (glycine N-choloyltransferase)
CBR1: Carbonyl reductase 1
CBR3: Carbonyl reductase 3
CBR4: Carbonyl reductase 4
CCBL1: Cysteine conjugate-beta lyase, cytoplasmic
CDA: Cytidine deaminase
CEL: Carboxyl ester lipase

CES1:	Carboxylesterase 1
CES1P1:	Carboxylesterase 1 pseudogene 1
CES2:	Carboxylesterase 2
CES3:	Carboxylesterase 3
CES5A:	Carboxylesterase 5A
CHST1:	Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1
CHST2:	Carbohydrate (N-acetylglucosamine-6-O) sulfotransferase 2
CHST3:	Carbohydrate (chondroitin 6) sulfotransferase 3
CHST4:	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 4
CHST5:	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 5
CHST6:	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 6
CHST7:	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 7
CHST8:	Carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 8
CHST9:	Carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 9
CHST10:	Carbohydrate sulfotransferase 10
CHST11:	Carbohydrate (chondroitin 4) sulfotransferase 11
CHST12:	Carbohydrate (chondroitin 4) sulfotransferase 12
CHST13:	Carbohydrate (chondroitin 4) sulfotransferase 13
COMT:	Catechol-O-methyltransferase
CYB5R3:	Cytochrome b5 reductase 3
CYP1A1:	Cytochrome P450, family 1, subfamily A, polypeptide 1
CYP1A2:	Cytochrome P450, family 1, subfamily A, polypeptide 2
CYP1B1:	Cytochrome P450, family 1, subfamily B, polypeptide 1
CYP2A6:	Cytochrome P450, family 2, subfamily A, polypeptide 6
CYP2A7:	Cytochrome P450, family 2, subfamily A, polypeptide 7
CYP2A13:	Cytochrome P450, family 2, subfamily A, polypeptide 13
CYP2B6:	Cytochrome P450, family 2, subfamily B, polypeptide 6
CYP2C8:	Cytochrome P450, family 2, subfamily C, polypeptide 8
CYP2C9:	Cytochrome P450, family 2, subfamily C, polypeptide 9
CYP2C18:	Cytochrome P450, family 2, subfamily C, polypeptide 18
CYP2C19:	Cytochrome P450, family 2, subfamily C, polypeptide 19
CYP2D6:	Cytochrome P450, family 2, subfamily D, polypeptide 6
CYP2D7P1:	Cytochrome P450, family 2, subfamily D, polypeptide 7 pseudogene 1
CYP2E1:	Cytochrome P450, family 2, subfamily E, polypeptide 1
CYP2F1:	Cytochrome P450, family 2, subfamily F, polypeptide 1
CYP2J2:	Cytochrome P450, family 2, subfamily J, polypeptide 2
CYP2R1:	Cytochrome P450, family 2, subfamily R, polypeptide 1
CYP2S1:	Cytochrome P450, family 2, subfamily S, polypeptide 1
CYP2W1:	Cytochrome P450, family 2, subfamily W, polypeptide 1
CYP3A4:	Cytochrome P450, family 3, subfamily A, polypeptide 4
CYP3A5:	Cytochrome P450, family 3, subfamily A, polypeptide 5
CYP3A7:	Cytochrome P450, family 3, subfamily A, polypeptide 7
CYP3A43:	Cytochrome P450, family 3, subfamily A, polypeptide 43
CYP4A11:	Cytochrome P450, family 4, subfamily A, polypeptide 11
CYP4A22:	Cytochrome P450, family 4, subfamily A, polypeptide 22
CYP4B1:	Cytochrome P450, family 4, subfamily B, polypeptide 1
CYP4F2:	Cytochrome P450, family 4, subfamily F, polypeptide 2
CYP4F3:	Cytochrome P450, family 4, subfamily F, polypeptide 3
CYP4F8:	Cytochrome P450, family 4, subfamily F, polypeptide 8
CYP4F11:	Cytochrome P450, family 4, subfamily F, polypeptide 11
CYP4F12:	Cytochrome P450, family 4, subfamily F, polypeptide 12
CYP4Z1:	Cytochrome P450, family 4, subfamily Z, polypeptide 1
CYP7A1:	Cytochrome P450, family 7, subfamily A, polypeptide 1
CYP7B1:	Cytochrome P450, family 7, subfamily B, polypeptide 1
CYP8B1:	Cytochrome P450, family 8, subfamily B, polypeptide 1
CYP11A1:	Cytochrome P450, family 11, subfamily A, polypeptide 1
CYP11B1:	Cytochrome P450, family 11, subfamily B, polypeptide 1
CYP11B2:	Cytochrome P450, family 11, subfamily B, polypeptide 2

CYP17A1: Cytochrome P450, family 17, subfamily A, polypeptide 1
CYP19A1: Cytochrome P450, family 19, subfamily A, polypeptide 1
CYP20A1: Cytochrome P450, family 20, subfamily A, polypeptide 1
CYP21A2: Cytochrome P450, family 21, subfamily A, polypeptide 2
CYP24A1: Cytochrome P450, family 24, subfamily A, polypeptide 1
CYP26A1: Cytochrome P450, family 26, subfamily A, polypeptide 1
CYP26B1: Cytochrome P450, family 26, subfamily B, polypeptide 1
CYP26C1: Cytochrome P450, family 26, subfamily C, polypeptide 1
CYP27A1: Cytochrome P450, family 27, subfamily A, polypeptide 1
CYP27B1: Cytochrome P450, family 27, subfamily B, polypeptide 1
CYP39A1: Cytochrome P450, family 39, subfamily A, polypeptide 1
CYP46A1: Cytochrome P450, family 46, subfamily A, polypeptide 1
CYP51A1: Cytochrome P450, family 51, subfamily A, polypeptide 1
DDOST: Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit (non-catalytic)
DHRS1: Dehydrogenase/reductase (SDR family) member 1
DHRS2: Dehydrogenase/reductase (SDR family) member 2
DHRS3: Dehydrogenase/reductase (SDR family) member 3
DHRS4: Dehydrogenase/reductase (SDR family) member 4
DHRS7: Dehydrogenase/reductase (SDR family) member 7
DHRS9: Dehydrogenase/reductase (SDR family) member 9
DHRS12: Dehydrogenase/reductase (SDR family) member 12
DHRS13: Dehydrogenase/reductase (SDR family) member 13
DHRSX: Dehydrogenase/reductase (SDR family) X-linked
DPEP1: Dipeptidase 1 (renal)
DPYD: Dihydropyrimidine dehydrogenase
EPHX1: Epoxide hydrolase 1, microsomal (xenobiotic)
EPHX2: Epoxide hydrolase 2, microsomal (xenobiotic)
ESD: Esterase D
FMO1: Flavin containing monooxygenase 1
FMO2: Flavin containing monooxygenase 2
FMO3: Flavin containing monooxygenase 3
FMO4: Flavin containing monooxygenase 4
FMO5: Flavin containing monooxygenase 5
FMO6P: Flavin containing monooxygenase 6 pseudogene
GAL3ST1: Galactose-3-O-sulfotransferase 1
GAMT: Guanidinoacetate N-methyltransferase
GLRX: Glutaredoxin (thioltransferase)
GLYAT: Glycine-N-acyltransferase
GNMT: Glycine N-methyltransferase
GPX1: Glutathione peroxidase 1
GPX2: Glutathione peroxidase 2 (gastrointestinal)
GPX3: Glutathione peroxidase 3 (plasma)
GPX4: Glutathione peroxidase 4
GPX5: Glutathione peroxidase 5
GPX6: Glutathione peroxidase 6 (olfactory)
GPX7: Glutathione peroxidase 7
GSR: Glutathione reductase
GSTA1: Glutathione S-transferase alpha 1
GSTA2: Glutathione S-transferase alpha 2
GSTA3: Glutathione S-transferase alpha 3
GSTA4: Glutathione S-transferase alpha 4
GSTA5: Glutathione S-transferase alpha 5
GSTCD: Glutathione S-transferase, C-terminal domain containing
GSTK1: Glutathione S-transferase kappa 1
GSTM1: Glutathione S-transferase mu 1
GSTM2: Glutathione S-transferase mu 2 (muscle)
GSTM3: Glutathione S-transferase mu 3 (brain)
GSTM4: Glutathione S-transferase mu 4
GSTM5: Glutathione S-transferase mu 5
GSTO1: Glutathione S-transferase omega 1
GSTO2: Glutathione S-transferase omega 2
GSTP1: Glutathione S-transferase pi 1
GSTT1: Glutathione S-transferase theta 1
GSTT2: Glutathione S-transferase theta 2
GSTZ1: Glutathione S-transferase zeta 1
GZMA: Granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3)
GZMB: Granzyme B (granzyme 2, cytotoxic T-lymphocyte-associated serine esterase 1)
HNMT: Histamine N-methyltransferase
HSD11B1: Hydroxysteroid (11-beta) dehydrogenase 1
HSD17B10: Hydroxysteroid (17-beta) dehydrogenase 10
HSD17B11: Hydroxysteroid (17-beta) dehydrogenase 11
HSD17B14: Hydroxysteroid (17-beta) dehydrogenase 14
INMT: Indolethylamine N-methyltransferase
MAOA: Monoamine oxidase A
MAOB: monoamine oxidase B
METAP1: Methionyl aminopeptidase 1
MGST1: Microsomal glutathione S-transferase 1
MGST2: Microsomal glutathione S-transferase 1
MGST3: Microsomal glutathione S-transferase 3
NAA20: N(alpha)-acetyltransferase 20, NatB catalytic subunit

NAT1:	N-acetyltransferase 1 (arylamine N-acetyltransferase)
NAT2:	N-acetyltransferase 2 (arylamine N-acetyltransferase)
NNMT:	Nicotinamide N-methyltransferase
NQO1:	NAD(P)H dehydrogenase, quinone 1
NQO2:	NAD(P)H dehydrogenase, quinone 2
PNMT:	Phenylethanolamine N-methyltransferase
PON1:	Paraoxonase 1
PON2:	Paraoxonase 2
PON3:	Paraoxonase 3
POR:	P450 (cytochrome) oxidoreductase
PTGES:	Prostaglandin E synthase
PTGS1:	Prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)
PTGS2:	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)
SAT1:	Spermidine/spermine N1-acetyltransferase 1
SMOX:	Spermine oxidase
SOD1:	Superoxide dismutase 1, soluble
SOD2:	Superoxide dismutase 2, mitochondrial
SULT1A1:	Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1
SULT1A2:	Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 2
SULT1A3:	Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 3
SULT1B1:	Sulfotransferase family, cytosolic, 1B, member 1
SULT1C1:	Sulfotransferase family, cytosolic, 1C, member 1
SULT1C2:	Sulfotransferase family, cytosolic, 1C, member 2
SULT1C3:	Sulfotransferase family, cytosolic, 1C, member 3
SULT1C4:	Sulfotransferase family, cytosolic, 1C, member 4
SULT1E1:	Sulfotransferase family 1E, estrogen-preferring, member 1
SULT2A1:	Sulfotransferase family, cytosolic, 2A, dehydroepiandrosterone (DHEA)-preferring, member 1
SULT2B1:	Sulfotransferase family, cytosolic, 2B, member 1
SULT4A1:	Sulfotransferase family 4A, member 1
SULT6B1:	sulfotransferase family, cytosolic, 6B, member 1
TBXAS1:	Thromboxane A synthase 1 (platelet)
TPMT:	Thiopurine S-methyltransferase
TST:	Thiopurine S-methyltransferase
UCHL1:	Ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)
UCHL3:	Ubiquitin carboxyl-terminal esterase L3 (ubiquitin thiolesterase)
UGT1A1:	UDP glucuronosyltransferase 1 family, polypeptide A1
UGT1A3:	UDP glucuronosyltransferase 1 family, polypeptide A3
UGT1A4:	UDP glucuronosyltransferase 1 family, polypeptide A4
UGT1A5:	UDP glucuronosyltransferase 1 family, polypeptide A5
UGT1A6:	UDP glucuronosyltransferase 1 family, polypeptide A6
UGT1A7:	UDP glucuronosyltransferase 1 family, polypeptide A7
UGT1A8:	UDP glucuronosyltransferase 1 family, polypeptide A8
UGT1A9:	UDP glucuronosyltransferase 1 family, polypeptide A9
UGT1A10:	UDP glucuronosyltransferase 1 family, polypeptide A10
UGT2A1:	UDP glucuronosyltransferase 2 family, polypeptide A1, complex locus
UGT2A3:	UDP glucuronosyltransferase 2 family, polypeptide A3
UGT2B10:	UDP glucuronosyltransferase 2 family, polypeptide B10
UGT2B11:	UDP glucuronosyltransferase 2 family, polypeptide B11
UGT2B15:	UDP glucuronosyltransferase 2 family, polypeptide B15
UGT2B17:	UDP glucuronosyltransferase 2 family, polypeptide B17
UGT2B28:	UDP glucuronosyltransferase 2 family, polypeptide B28
UGT2B4:	UDP glucuronosyltransferase 2 family, polypeptide B4
UGT2B7:	UDP glucuronosyltransferase 2 family, polypeptide B7
UGT3A1:	UDP glycosyltransferase 3 family, polypeptide A1
UGT8:	UDP glycosyltransferase 8
XDH:	Xanthine dehydrogenase

REFERENCES

[1] National Center for Biotechnology Information. The NCBI database. http://www.ncbi.nlm.nih.gov/pubmed/; 2013 [accessed 19.06.13].

[2] Suehs BT, Davis CD, Alvir J, van Amerongen D, Pharmd NC, Joshi AV, et al. The clinical and economic burden of newly diagnosed Alzheimer's disease in a medicare advantage population. Am J Alzheimers Dis Other Demen 2013;28(4):384–92.

[3] National Center for Health Statistics. Health, United States, 2009: With special feature on medical technology. Hyattsville, MD. 2010.

[4] Centers for Disease Control and Prevention. <http://www.cdc.gov/DataStatistics/>. Updated March 6 2012 [accessed 19.06.13].

[5] Sousa RM, Ferri CP, Acosta D, Albanese E, Guerra M, Huang Y, et al. Contribution of chronic diseases to disability in elderly people in countries with low and middle incomes: a 10/66 Dementia Research Group population-based survey. Lancet 2009;374(9704):1821–30.

[6] Herrmann N, Tam DY, Balshaw R, Sambrook R, Lesnikova N, Lanctôt KL, et al. The relation between disease severity and cost of caring for patients with Alzheimer disease in Canada. Can J Psychiatry 2010;55(12):768–75.

[7] Cacabelos R. The path to personalized medicine in mental disorders. In: Ritsner MS, editor. The handbook of neuropsychiatric biomarkers, endophenotypes and genes, Vol. 4. Netherlands: Springer; 2009. p. 3–63.

[8] Wimo A, Jönsson L, Gustavsson A, McDaid D, Ersek K, Georges J, et al. The economic impact of dementia in Europe in 2008-cost estimates from the Eurocode project. Int J Geriatr Psychiatry 2011;26(8):825–32.

[9] Wimo A, Reed CC, Dodel R, Belger M, Jones RW, Happich M, et al. The GERAS study: a prospective observational study of costs and resource use in community dwellers with Alzheimer's disease in three European countries: study design and baseline findings. J Alzheimers Dis 2013;36(2):385–99.

[10] Kamble P, Chen H, Sherer JT, Aparasu RR. Use of antipsychotics among elderly nursing home residents with dementia in the US: an analysis of National Survey Data. Drugs Aging 2009;26(6):483–92.

[11] Liperoti R, Onder G, Landi F, Lapane KL, Mor V, Bernabei R, et al. All-cause mortality associated with atypical and conventional antipsychotics among nursing home residents with dementia: a retrospective cohort study. J Clin Psychiatry 2009;70(10):1340–7.

[12] Cacabelos R. Pharmacogenomics and therapeutic strategies for dementia. Expert Rev Mol Diagn 2009;9(6):567–611.

[13] Anderson CNG, Grant SGN. High throughput protein expression screening in the nervous system-needs and limitations. J Physiol 2006;575(Pt 2):367–72.

[14] Xu X, Zhan M, Duan W, Prabhu V, Brenneman R, Wood W, et al. Gene expression atlas of the mouse central nervous system: impact and interactions of age, energy intake and gender. Genome Biol 2007;8(11):R234.

[15] Cacabelos R, Fernández-Novoa L, Martínez-Bouza R, McKay A, Carril JC, Lombardi V, et al. Future trends in the pharmacogenomics of brain disorders and dementia: influence of APOE and CYP2D6 variants. Pharmaceuticals 2010;3(10):3040–100.

[16] Cacabelos R, Fernández-Novoa L, Lombardi V, Kubota Y, Takeda M. Molecular genetics of Alzheimer's disease and aging. Methods Find Exp Clin Pharmacols 2005;27(Suppl. A):1–573.

[17] Alzheimer Research Forum. Alzforum. <http://www.alzgene.org/> Published 1996. Updated April 18 2011 [accessed 19.06.13].

[18] Cacabelos R. Alzheimer's disease 2011: where are we heading? Gen-T 2011;8:54–86.

[19] Selkoe DJ, Podlisny MB. Deciphering the genetic basis of Alzheimer's disease. Annu Rev Genomics Hum Genet 2002;3:67–99.

[20] Suh YH, Checler F. Amyloid precursor protein, presenilins, and α-synuclein: molecular pathogenesis and pharmacological applications in Alzheimer's disease. Pharmacol Rev 2002;54(3):469–525.

[21] Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science 2002;297(5580):353–6.

[22] Takeda M, Martínez R, Kudo T, Tanaka T, Okochi M, Tagami S, et al. Apolipoprotein E and central nervous system disorders: reviews of clinical findings. Psychiatry Clin Neurosci 2010;64(6):592–607.

[23] Cacabelos R. Pharmacogenomics in Alzheimer's disease. Methods Mol Biol 2008;448:213–357.

[24] Cacabelos R, Takeda M. Pharmacogenomics, nutrigenomics and future therapeutics in Alzheimer's disease. Drugs Future 2006;31(Suppl. B):5–146.

[25] Cacabelos R. The application of functional genomics to Alzheimer's disease. Pharmacogenomics 2003;4(5):597–621.

[26] Petrlova J, Hong HS, Bricarello DA, Harishchandra G, Lorigan GA, Jin LW, et al. A differential association of apolipoprotein E isoforms with the amyloid-β oligomer in solution. Proteins 2011;79(2):402–16.

[27] Samaranch L, Cervantes S, Barabash A, Alonso A, Cabranes JA, Lamet I, et al. The effect of MAPT H1 and APOE ε4 on transition from mild cognitive impairment to dementia. J Alzheimers Dis 2010;22(4):1065–71.

[28] Cacabelos R. Donepezil in Alzheimer's disease: from conventional trials to pharmacogenetics. Neuropsychiatr Dis Treat 2007;3(3):303–33.

[29] Mallick B, Ghosh Z. A complex crosstalk between polymorphic microRNA target sites and AD prognosis. RNA Biol 2011;8(4):665–73.

[30] Qureshi IA, Mehler MF. Advances in epigenetics and epigenomics for neurodegenerative diseases. Curr Neurol Neurosci Rep 2011;11(5):464–73.

[31] Enciu AM, Popescu BO, Gheorghisan-Galateanu A. MicroRNAs in brain development and degeneration. Mol Biol Rep 2012;39(3):2243–52.

[32] Murray IV, Proza JF, Sohrabji F, Lawler JM. Vascular and metabolic dysfunction in Alzheimer's disease: a review. Exp Biol Med (Maywood) 2011;236(7):772–82.

[33] Ettorre E, Cerra E, Marigliano B, Vigliotta M, Vulcano A, Fossati C, et al. Role of cardiovascular risk factors (CRF) in the patients with mild cognitive impairment (MCI). Arch Gerontol Geriatr 2012;54(2):330–2.

[34] Lin KP, Chen SY, Lai LC, Huang YL, Chen JH, Chen TF, et al. Genetic polymorphisms of a novel vascular susceptibility gene, Ninjurin2 (NINJ2), are associated with a decreased risk of Alzheimer's disease. PLoS One 2011;6(6):e20573.

[35] Romero A, Cacabelos R, Oset-Gasque MJ, Samadi A, Marco-Contelles J. Novel tacrine-related drugs as potential candidates for the treatment of Alzheimer's disease. Bioorg Med Chem Lett 2013;23(7):1916–22.

[36] Aliev G, Palacios HH, Cacabelos P, Cacabelos R, Burzynski G, Burzynski SR. Mitochondria specific antioxidants and their derivatives in the context of the drug development for neurodegeneration and cancer. Drug Des 2013;2:1.

[37] Grammas P, Martínez J, Miller B. Cerebral microvascular endothelium and the pathogenesis of neurodegenerative diseases. Expert Rev Mol Med 2011;13:e19.

[38] Cacabelos R, Fernández-Novoa L, Lombardi V, Corzo L, Pichel V, Kubota Y. Cerebrovascular risk factors in Alzheimer's disease: brain hemodynamics and pharmacogenomic implications. Neurol Res 2003;25(6):567–80.

[39] Cacabelos R, Fernández-Novoa L, Corzo L, Amado L, Pichel V, Lombardi V, et al. Phenotypic profiles and functional genomics in Alzheimer's disease and in dementia with a vascular component. Neurol Res 2004;26(5):459–80.

[40] Kim JH, Hwang KJ, Kim JH, Lee YH, Rhee HY, Park KC. Regional white matter hyperintensities in normal aging, single domain

amnestic mild cognitive impairment, and mild Alzheimer's disease. J Clin Neurosci 2011;18(8):1101–6.

[41] Chen H, Zhang JH. Cerebral amyloid angiopathy-related microhemorrhages in Alzheimer's disease: a review of investigative animal models. Acta Neurochir Suppl 2011;111:15–17.

[42] Yates PA, Sirisriro R, Villemagne VL, Farquharson S, Masters CL, Rowe CCAIBL Research Group. Cerebral microhemorrhage and brain β-amyloid in aging and Alzheimer disease. Neurology 2011;77(1):48–54.

[43] Brenn A, Grube M, Peters M, Fischer A, Jedlitschky G, Kroemer HK, et al. Beta-amyloid downregulates MDR1-P-glycoprotein (Abcb1) expression at the blood-brain barrier in mice. Int J Alzheimers Dis 2011;2011:690121.

[44] Chen KH, Reese EA, Kim HW, Rapoport SI, Rao JS. Disturbed neurotransmitter transporter expression in Alzheimer's disease brain. J Alzheimers Dis 2011;26(4):755–66.

[45] Aliev G, Obrenovich ME, Tabrez S, Jabir NR, Reddy VP, Li Y, et al. Link between cancer and Alzheimer disease via oxidative stress induced by nitric oxide-dependent mitochondrial DNA overproliferation and deletion. Oxid Med Cell Longev 2013;2013:962984.

[46] Tayler H, Fraser T, Miners JS, Kehoe PG, Love S. Oxidative balance in Alzheimer's disease: relationship to APOE, Braak tangle stage, and the concentrations of soluble and insoluble amyloid-β. J Alzheimers Dis 2010;22(4):1363–73.

[47] Alikhani N, Guo L, Yan S, Du H, Pinho CM, Chen JX, et al. Decreased proteolytic activity of the mitochondrial amyloid-β degrading enzyme, PreP peptidasome, in Alzheimer's disease brain mitochondria. J Alzheimers Dis 2011;27(1):75–87.

[48] Mathew A, Yoshida Y, Maekawa T, Sakthi Kumar DS. Alzheimer's disease: cholesterol a menace? Brain Res Bull 2011;86(1–2):1–12.

[49] Qiu L, Buie C, Reay A, Vaughn MW, Cheng KH. Molecular dynamics simulations reveal the protective role of cholesterol in beta amyloid protein-induced membrane disruptions in neuronal membrane mimics. J Phys Chem B 2011;115(32):9795–812.

[50] Jones L, Holmans PA, Hamshere ML, Harold D, Moskvina V, Ivanov D, et al. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. PLoS One 2010;5(11):e13950.

[51] Pan XD, Zhu YG, Lin N, Zhang J, Ye QY, Huang HP, et al. Microglial phagocytosis induced by fibrillar beta-amyloid is attenuated by oligomeric beta-amyloid: implications for Alzheimer's disease. Mol Neurodegener 2011;6:45.

[52] Diniz BS, Teixeira AL, Ojopi EB, Talib LL, Mendonça VA, Gattaz WF, et al. Higher serum sTNFR1 level predicts conversion from mild cognitive impairment to Alzheimer's disease. J Alzheimers Dis 2010;22(4):1305–11.

[53] Cacabelos R, Martínez-Bouza R. Genomics and pharmacogenomics of dementia. CNS Neurosci Ther 2011;17(5):566–76.

[54] Cacabelos R. Pharmacogenomics and therapeutic prospects in Alzheimer's disease. Expert Opin Pharmacother 2005;6(12):1967–87.

[55] Cacabelos R. Pharmacogenomics and therapeutic prospects in dementia. Eur Arch Psychiatry Clin Neurosci 2008;258(Suppl. 1):28–47.

[56] Cacabelos R. Pharmacogenetic basis for therapeutic optimization in Alzheimer's disease. Mol Diagn Ther 2007;11(6):385–405.

[57] Cacabelos R, Llovo R, Fraile C, Fernández-Novoa L. Pharmacogenetic aspects of therapy with cholinesterase inhibitors: the role of CYP2D6 in Alzheimer's disease pharmacogenetics. Curr Alzheimer Res 2007;4(4):479–500.

[58] Cacabelos R. Molecular pathology and pharmacogenomics in Alzheimer's disease: polygenic-related effects of multifactorial treatments on cognition, anxiety, and depression. Methods Find Exp Clin Pharmacol 2007;29(Suppl. A):1–91.

[59] Cacabelos R. Pharmacogenomics, nutrigenomics and therapeutic optimization in Alzheimer's disease. Aging Health 2005;1(2):303–48.

[60] Braak H, del Tredici K. The pathological process underlying Alzheimer's disease in individuals under thirty. Acta Neuropathol 2011;121(2):171–81.

[61] Lu PH, Thompson PM, Leow A, Lee GJ, Lee A, Yanovsky I, et al. Apolipoprotein E genotype is associated with temporal and hippocampal atrophy rates in healthy elderly adults: a tensor-based morphometry study. J Alzheimers Dis 2011;23(3):433–42.

[62] Canu E, Frisoni GB, Agosta F, Pievani M, Bonetti M, Filippi M. Early and late onset Alzheimer's disease patients have distinct patterns of white matter damage. Neurobiol Aging 2012;33(6):1023–33.

[63] Herholz K, Ebmeier K. Clinical amyloid imaging in Alzheimer's disease. Lancet Neurol 2011;10(7):667–70.

[64] Vialatte FB, Dauwels J, Maurice M, Musha T, Cichocki A. Improving the specificity of EEG for diagnosing Alzheimer's disease. Int J Alzheimers Dis 2011;2011:259069.

[65] Roh JH, Park MH, Ko D, Park KW, Lee DH, Han C, et al. Region and frequency specific changes of spectral power in Alzheimer's disease and mild cognitive impairment. Clin Neurophysiol 2011;122(11):2169–76.

[66] Canuet L, Tellado I, Couceiro V, Fraile C, Fernandez-Novoa L, Ishii R, et al. Resting-state network disruption and APOE genotype in Alzheimer's disease: a lagged functional connectivity study. PLoS One 2012;7(9):e46289.

[67] Cacabelos R, Fernández-Novoa Corzo L, Pichel V, Lombardi V, Kubota Y. Genomics and phenotypic profiles in dementia: implications for pharmacological treatment. Methods Find Exp Clin Pharmacol 2004;26(6):421–44.

[68] Hampel H, Frank R, Broich K, Teipel SJ, Katz RG, Hardy J, et al. Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. Nat Rev Drug Discov 2010;9(7):560–74.

[69] Cacabelos R, Corzo L, Fernández-Novoa L, Lombardi V. Histamine in Alzheimer's disease pathogenesis: biochemistry and functional genomics. Methods Find Exp Clin Pharmacol 2004;26(Suppl. 2):9–16.

[70] Portelius E, Mattsson N, Andreasson U, Blennow K, Zetterberg H. Novel Aβ isoforms in Alzheimer's disease – their role in diagnosis and treatment. Curr Pharm Des 2011;17(25):2594–602.

[71] Kester MI, Scheffer PG, Koel-Simmelink MJ, Twaalfhoven H, Verwey NA, Veerhuis R, et al. Serial CSF sampling in Alzheimer's disease: specific versus non-specific markers. Neurobiol Aging 2012;33(8):1591–8.

[72] Roher AE, Maarouf CL, Sue LI, Hu Y, Wilson J, Beach TG. Proteomics-derived cerebrospinal fluid markers of autopsy-confirmed Alzheimer's disease. Biomarkers 2009;14(7):493–501.

[73] Popp J, Lewczuk P, Frommann I, Kölsch H, Kornhuber J, Maier W, et al. Cerebrospinal fluid markers for Alzheimer's disease over the lifespan: effects of age and the APOEε4 genotype. J Alzheimers Dis 2010;22(2):459–68.

[74] Han MR, Schellenberg GD, Wang LS. Alzheimer's disease neuroimaging initiative. Genome-wide association reveals genetic effects

on human Aβ42 and τ protein levels in cerebrospinal fluids: a case control study. BMC Neurol 2010;10:90.

[75] Thambisetty M, Tripaldi R, Riddoch-Contreras J, Hye A, An Y, Campbell J, et al. Proteome-based plasma markers of brain amyloid-β deposition in non-demented older individuals. J. Alzheimers Dis 2010;22(4):1099–109.

[76] Loveman E, Green C, Kirby J, Takeda A, Picot J, Payne E, et al. The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease. Health Technol Assess 2006;10(1):1–160.

[77] Cacabelos R, Álvarez A, Lombardi V, Fernández-Novoa L, Corzo L, Pérez P, et al. Pharmacological treatment of Alzheimer disease: from psychotropic drugs and cholinesterase inhibitors to pharmacogenomics. Drugs Today (Barc) 2000;36(7):415–99.

[78] Giacobini E. The brain cholinergic system in health and disease. In: Giacobini E, Pepeu G, editors. Cholinesterases in human brain: the effect of cholinesterase inhibitors on Alzheimer's disease and related disordersp. Oxon, UK: Informa Healthcare; 2006. p. 235–64. [chapter 18]

[79] Reisberg B, Doody R, Stoffler A, Schmitt F, Ferris S, Möbius HJ. Memantine in moderate-to-severe Alzheimer's disease. N Engl J Med 2003;348(14):1333–41.

[80] Schenk DB, Seubert P, Grundman M, Black R. Aβ immunotherapy: lessons learned for potential treatment of Alzheimer's disease. Neurodegener Dis 2005;2(5):255–60.

[81] Wisniewski T, Boutajangout A. Vaccination as a therapeutic approach to Alzheimer's disease. Mt Sinai J Med 2010;77(1):17–31.

[82] de Strooper B, Vassar R, Golde T. The secretases: enzymes with therapeutic potential in Alzheimer disease. Nat Rev Neurol 2010;6(2):99–107.

[83] Shelton CC, Zhu L, Chau D, Yang L, Wang R, Djaballah H, et al. Modulation of gamma-secretase specificity using small molecule allosteric inhibitors. Proc Natl Acad Sci USA 2009;106(48):20228–33.

[84] Lambracht-Washington D, Qu BX, Fu M, Eagar TN, Stüve O, Rosenberg RN. DNA beta-amyloid(1–42) trimer immunization for Alzheimer disease in a wild-type mouse mode. JAMA 2009;302(16):1796–802.

[85] Carrera I, Etcheverría I, Fernández-Novoa L, Lombardi L, Vigo C, Cacabelos R. Development of the EuroEspes EB101 vaccine against Alzheimer's disease. Gen-T 2011;8:38–52.

[86] Wisniewski T, Konietzko U. Amyloid-beta immunisation for Alzheimer's disease. Lancet Neurol 2008;7(9):805–11.

[87] Singh S, Kushwah AS, Singh R, Farswan M, Kaur R. Current therapeutic strategy in Alzheimer's disease. Eur Rev Med Pharmacol Sci 2012;16(12):1651–64.

[88] Grill JD, Cummings JL. Current therapeutic targets for the treatment of Alzheimer's disease. Expert Rev Neurother 2010;10(5):711–28.

[89] Tabira T. Immunization therapy for Alzheimer disease: a comprehensive review of active immunization strategies. Tohoku J Exp Med 2010;220(2):95–106.

[90] Carrera I, Cacabelos R. Novel immunotherapeutic procedures for prevention of Alzheimer's disease. Drug Des 2013;2:107.

[91] Galimberti D, Ghezzi L, Scarpini E. Immunotherapy against amyloid pathology in Alzheimer's disease. J Neurol Sci 2013;333(1–2):50–4.

[92] Maeda S, Sahara N, Saito Y, Murayama S, Ikai A, Takashima A. Increased levels of granular tau oligomers: an early sign of brain aging and Alzheimer's disease. Neurosci Res 2006;54(3):197–201.

[93] Boutajangout A, Sigurdsson EM, Krishnamurthy PK. Tau as a therapeutic target for Alzheimer's disease. Curr Alzheimer Res 2011;8(6):666–77.

[94] Johnson GV. Tau phosphorylation and proteolysis: insights and perspectives. J Alzheimers Dis 2006;9(Suppl. 3):243–50.

[95] Kudo LC, Parfenova L, Ren G, Vi N, Hui M, Ma Z, et al. Puromycin-sensitive aminopeptidase (PSA/NPEPPS) impedes development of neuropathology in hPSA/TAU(P301L) double-transgenic mice. Hum Mol Genet 2011;20(9):1820–33.

[96] Trojanowski JQ, Smith AB, Huryn D, Lee VM. Microtubule-stabilising drugs for therapy of Alzheimer's disease and other neurodegenerative disorders with axonal transport impairments. Expert Opin Pharmacother 2005;6(5):683–6.

[97] Zhang B, Maiti A, Shively S, Lakhani F, McDonald-Jones G, Bruce J, et al. Microtubule-binding drugs offset tau sequestration by stabilizing microtubules and reversing fast axonal transport deficits in a tauopathy model. Proc Natl Acad Sci USA 2005;102(1):227–31.

[98] Boutajangout A, Quartermain D, Sigurdsson EM. Immunotherapy targeting pathological tau prevents cognitive decline in a new tangle mouse model. J Neurosci 2010;30(49):16559–66.

[99] Schenk D, Barbour R, Dunn W, Gordon G, Grajeda H, Guido T, et al. Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse. Nature 1999;400(6740):173–7.

[100] Janus C, Pearson J, McLaurin J, Mathews PM, Jiang Y, Chishti MA, et al. A beta peptide immunization reduces behavioural impairment and plaques in a model of Alzheimer's disease. Nature 2000;408(6815):979–82.

[101] Bacskai BJ, Kajdasz ST, Christie RH, Carter C, Games D, Seubert P, et al. Imaging of amyloid-beta deposits in brains of living mice permits direct observation of clearance of plaques with immunotherapy. Nat Med 2001;7(3):369–72.

[102] Imbimbo BP, Ottonello S, Frisardi V, Solfrizzi V, Greco A, Seripa D, et al. Solanezumab for the treatment of mild-to-moderate Alzheimer's disease. Expert Rev Clin Immunol 2012;8(2):135–49.

[103] Panza F, Frisardi V, Imbimbo BP, D'Onofrio G, Pietrarossa G, Seripa D, et al. Bapineuzumab: anti-β-amyloid monoclonal antibodies for the treatment of Alzheimer's disease. Immunotherapy 2010;2(6):767–82.

[104] Magga J, Puli L, Pihlaja R, Kanninen K, Neulamaa S, Malm T, et al. Human intravenous immunoglobulin provides protection against Aβ toxicity by multiple mechanisms in a mouse model of Alzheimer's disease. J Neuroinflammation 2010;7:90.

[105] Kingwell K. Alzheimer disease: amyloid-β immunotherapy CAD106 passes first safety test in patients with Alzheimer disease. Nat Rev Neurol 2012;8(8):414.

[106] Delrieu J, Ousset PJ, Caillaud C, Vellas B. 'Clinical trials in Alzheimer's disease': immunotherapy approaches. J Neurochem 2012;120(Suppl. 1):186–93.

[107] Qu BX, Xiang Q, Li L, Johnston SA, Hynan LS, Rosenberg RN. Abeta42 gene vaccine prevents Abeta42 deposition in brain of double transgenic mice. J Neurol Sci 2007;260(1–2):204–13.

[108] Roses AD. The medical and economic roles of pipeline pharmacogenetics: Alzheimer's disease as a model of efficacy and HLA-B(*)5701 as a model of safety. Neuropsychopharmacology 2009;34(1):6–17.

[109] Roses AD, Saunders AM, Huang Y, Strum J, Weisgraber KH, Mahley RW. Complex disease-associated pharmacogenetics: drug efficacy, drug safety, and confirmation of a pathogenic hypothesis (Alzheimer's disease). Pharmacogenomics J 2007;7(1):10–28.

[110] Risner ME, Saunders AM, Altman JF, Ormandy GC, Craft S, Foley IM, et al. Efficacy of rosiglitazone in a genetically defined population with mild-to-moderate Alzheimer's disease. Pharmacogenomics J 2006;6(4):246–54.
[111] Cacabelos R. Pharmacogenomics of central nervous system (CNS) drugs. Drug Dev Res 2012;73:461–76.
[112] Xie HG, Kim RB, Wood AJ, Stein CM. Molecular basis of ethnic differences in drug disposition and response. Annu Rev Pharm Toxicol 2001;41:815–50.
[113] Cacabelos R, editor. World guide for drug use and pharmacogenomics. Corunna, Spain: EuroEspes Publishing; 2012.
[114] Division of Pharmacology/Neurobiology, Biozentrum, University of Basel, Switzerland. The human cytochrome P450 (CYP) allele nomenclature database. <http://www.cypalleles.ki.se/>; 2013 [accessed 19.06.13].
[115] Stanford University. PharmGK. <http:// www.pharmgkb.org> Published 2001 [accessed 19.06.13].
[116] Isaza CA, Henao J, López AM, Cacabelos R. Isolation, sequence and genotyping of the drug metabolizer CYP2D6 gene in the Colombian population. Meth Find Exp Clin Pharmacol 2000;22(9):695–705.
[117] Mizutani T. PM frequencies of major CYPs in Asians and Caucasians. Drug Metab Rev 2003;35(2–3):99–106.
[118] Ozawa S, Soyama A, Saeki M, Fukushima-Uesaka H, Itoda M, Koyano S, et al. Ethnic differences in genetic polymorphisms of CYP2D6, CYP2C19, CYP3As and MDR1/ABCB1. Drug Metab Pharmacokin 2004;19(2):83–95.
[119] Weinshilboum RM, Wang L. Pharmacogenetics and pharmacogenomics: development, science, and translation. Annu Rev Genomics Hum Genet 2006;7:223–45.
[120] Marquez B, Van Bambeke F. ABC multidrug transporters: target for modulation of drug pharmacokinetics and drug-drug interactions. Curr Drug Targets 2011;12(5):600–20.
[121] Haufroid V. Genetic polymorphisms of ATP-binding cassette transporters ABCB1 and ABCC2 and their impact on drug disposition. Curr Drug Targets 2011;12(5):631–46.
[122] Cacabelos R. The metabolomics paradigm of pharmacogenomics in complex disorders. Metabolomics 2012;2:5.
[123] Hosoya K, Tachikawa M. Roles of organic anion/cation transporters at the blood-brain and blood-cerebrospinal fluid barriers involving uremic toxins. Clin Exp Nephrol 2011;15(4):478–85.
[124] Carl SM, Lindley DJ, Couraud PO, Weksler BB, Romero I, Mowery SA. Knipp GT.ABC and SLC transporter expression and Pot substrate characterization across the human CMEC/D3 blood-brain barrier cell line. Mol Pharm 2010;7(4):1057–68.
[125] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack Jr CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7(3):263–9.
[126] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7(3):270–9.
[127] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7(3):280–92.
[128] Goldman JS, Hahn SE, Catania JW, LaRusse-Eckert S, Butson MB, Rumbaugh M, et al. Genetic counseling and testing for Alzheimer disease: joint practice guidelines of the American College of Medical Genetics and the National Society of Genetic Counselors. Genet Med 2011;13(6):597–605.
[129] Schipper HM. Presymptomatic apolipoprotein E genotyping for Alzheimer's disease risk assessment and prevention. Alzheimers Dement 2011;7(4):e118–23.
[130] Kopits IM, Chen C, Roberts JS, Uhlmann W, Green RC. Willingness to pay for genetic testing for Alzheimer's disease: a measure of personal utility. Genet Test Mol Biomarkers 2011;15(12):871–5.
[131] Ohara T, Ninomiya T, Kubo M, Hirakawa Y, Doi Y, Hata J, et al. Apolipoprotein genotype for prediction of Alzheimer's disease in older Japanese: the hisayama study. J Am Geriatr Soc 2011;59(6):1074–9.
[132] Akinleye I, Roberts JS, Royal CD, Linnenbringer E, Obisesan TO, Fasaye GA, et al. Differences between African American and white research volunteers in their attitudes, beliefs and knowledge regarding genetic testing for Alzheimer's disease. J Genet Couns 2011;20(6):650–9.
[133] Sowell RA, Owen JB, Butterfield DA. Proteomics in animal models of Alzheimer's and Parkinson's disease. Ageing Res Rev 2009;8(1):1–17.
[134] Britschgi M, Rufibach K, Bauer Huang SL, Clark CM, Kaye JA, Li G, et al. Modeling of pathological traits in Alzheimer's disease based on systemic extracellular signaling proteome. Mol Cell Proteomics 2011;10(10):M111.008862.
[135] Khachaturian ZS, Petersen RC, Gauthier S, Buckholtz N, Corey-Bloom JP, Evans B, et al. A roadmap for the prevention of dementia: the inaugural Leon Thal Symposium. Alzheimers Dement 2008;4(3):156–63.
[136] Cacabelos R. Role of nutrition in the prevention of Alzheimer's disease. Aging Health 2005;1(3):359–62.
[137] Patel L, Grossberg GT. Combination therapy for Alzheimer's disease. Drugs Aging 2011;28(7):539–46.
[138] Álvarez XA, Cacabelos R, Sampedro C, Couceiro V, Aleixandre M, Vargas M, et al. Combination treatment in Alzheimer's disease: results of a randomized, controlled trial with cerebrolysin and donepezil. Curr Alzheimer Res 2011;8(5):583–91.
[139] Need AC, Motulsky AG, Goldstein DB. Priorities and standards in pharmacogenetic research. Nat Genet 2005;37(7):671–81.
[140] Carril JC, Martínez R, Fernández-Novoa L, Carrera I, Corzo L, Fraile C, et al. The EuroEspes pharmacogenetic card. Personalization in pharmacological treatment. Gen-T 2011;3:89–102.

This page is intentionally left blank

Chapter 28

Asthma

Dominick Shaw, Michael Portelli, and Ian Sayers
Division of Respiratory Medicine, University of Nottingham, UK

Chapter Outline

28.1 Overview **617**
28.1.1 Definition 618
28.1.2 Incidence/Prevalence 618
28.1.3 Cost 618
28.2 Clinical Aspects **618**
28.2.1 Symptoms 618
28.2.2 Associated Comorbidities and Confounders 619
28.2.3 Asthma Control 619
28.2.4 Exacerbations 620
28.2.5 Diagnosis 621
28.3 Asthma Pathophysiology **623**
28.3.1 Biological Basis 623
28.3.2 Genetic Basis 623
28.4 Current Management of Asthma **629**
28.4.1 Bronchodilators 629
28.4.2 Inhaled Corticosteroids 629
28.4.3 Other Pharmacological Treatments and Alternative Therapies 630
28.4.4 Specific Treatment for Severe Asthma 631
28.5 Stratified Medicine for Asthma **631**
28.5.1 Heterogeneity in Asthma 631
28.5.2 Problems with Current Therapy 632
28.5.3 Nongenetic Approaches to Patient Stratification 633
28.5.4 Clinicopathophysiological Measures in Asthma with Potential Use in Stratification 634
28.5.5 Current Asthma Therapies—Pharmacogenetic Approaches to Stratification 636
28.5.6 Pharmacogenetic Approaches to Disease Management: *ADRB2* Arg16 and Second-Line Therapy 640
28.5.7 Leukotriene Synthesis Inhibitor and Leukotriene Receptor Antagonist Pharmacogenetics 640
28.5.8 Corticosteroid Pharmacogenetics 643
28.6 Drug Development in Asthma and Emerging Targets **645**
28.6.1 Screening, Stratification, Personalization 646
28.7 Conclusions and Future Directions **647**
Acknowledgments **647**
References **647**

28.1 OVERVIEW

The term "*asthma*" originates from the Greek, meaning short of breath. The term was refined in the nineteenth century by Henry Hyde Salter (1823–1871) with the publication of his *magnum opus* "On Asthma and its Treatment," based on his experience of hundreds of cases. Salter defined asthma as "Paroxysmal dyspnea of a peculiar character with intervals of healthy respiration between attacks," which summarizes the episodic features and symptoms of the disease.

The use of the word *asthma* to define disease has recently attracted criticism [1], and it is now realized that asthma covers a host of overlapping, interrelated, and variable features, including abnormalities of airway inflammation, airway smooth-muscle function, airway immunology, and the airway epithelium [2], all of which contribute to the disease process and lead to symptoms of wheeze, cough, breathlessness, and chest tightness. These symptoms can change over time with or without treatment. Episodes of exacerbation (asthma attacks), which require treatment with systemic steroids can also occur and account for a large part of the disease burden.

A great deal of progress has been made in understanding asthma, with the identification of different clinical phenotypes (i.e., clinically relevant properties) and disease endotypes (pathobiological mechanisms behind disease features). Medicines applicable to an individual rather than to the population are being developed, and a stratified medicinal approach is being applied to both treatment design and disease management. Our understanding of asthma phenotypes remains incomplete because of a lack of large longitudinal studies, the difficulties caused by the impact of treatment on the results of unbiased approaches to disease stratification, a lack of a clear disease definition, and clinical investigations that have poor diagnostic utility. Understanding asthma and the diseases present under this

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00028-1
© 2014 Elsevier Inc. All rights reserved.

umbrella definition will require integrating bench and bedside approaches using data from ongoing genomic and proteomic profiling studies of large, well-characterized asthma populations—such as the current EU-wide UBIOPRED study (formally, "Unbiased Biomarkers in Prediction of Respiratory Disease Outcomes")—with feedback to improved *in vitro* and *in vivo* animal models and human studies.

28.1.1 Definition

The *International Consensus Report on the Diagnosis and Treatment of Asthma* defines asthma as "a chronic inflammatory disorder of the airways in which many cells play a role, including mast cells and eosinophils. In susceptible individuals this inflammation causes symptoms which are usually associated with widespread, but variable, airflow obstruction that is often reversible either spontaneously or with treatment, and causes an associated increase in airway responsiveness to a variety of stimuli." The interaction of these features determines the clinical manifestations and severity of the disease and the response to treatment [3]. How these features relate to each other, how they are best measured, and how they contribute to clinical manifestations of the disease remains unclear. Age of disease onset also complicates understanding, although there are many shared features in the diagnosis of asthma in children and in adults, there are also important differences. The differential diagnosis, the natural history of wheezing illnesses, and the ability to perform certain investigations are all influenced by age [4].

28.1.2 Incidence/Prevalence

According to the World Health Organization (WHO), between 100 and 150 million people around the globe (roughly equivalent to the population of the Russian Federation) suffer from asthma. Worldwide, deaths from asthma have reached more than 180,000 annually. Overall, asthma affects 5–10% of the population in many developed countries. One study found that the annual estimated incidence of physician-diagnosed asthma ranged from 0.6–29.5 per 1000 persons. Risk factors for incident asthma among children include male gender, atopic sensitization, parental history of asthma, early-life stressors and infections, obesity, and exposure to indoor allergens, tobacco smoke, and outdoor pollutants. Risk factors for adult-onset asthma include female sex, airway hyper-responsiveness, lifestyle factors, and work-related exposures.

28.1.3 Cost

While the exact cost of asthma care is difficult to determine, a 2009 systematic review [5] found eight national studies reporting a national total cost. The total costs in 2008 in U.S. dollars were high and varied widely: Singapore, $49.36M; Canada, $654M; Switzerland, $1413M; Germany, $2740M; United States, $8256M. Although the costs of asthma care vary by country, it is estimated that worldwide the number of disability-adjusted life years (DALYs) lost from asthma is about 15 million per year. Worldwide asthma accounts for around 1% of all DALYs lost, which reflects the high prevalence and severity of asthma. The number of asthma-caused lost DALYs is similar to that for diabetes, cirrhosis of the liver, and schizophrenia. Costs are related to other factors, including asthma control, demographics, medical history, and doses of inhaled corticosteroid (ICS) prescribed [6]. Improving asthma control is associated with decreased cost.

28.2 CLINICAL ASPECTS

28.2.1 Symptoms

Asthma classically causes wheeze, cough, chest tightness, and breathlessness (dyspnoea). The pattern of symptoms can be a clue to the diagnosis. Episodic symptoms worse at night and in the early morning, or in response to exercise, allergen and cold air exposure, or after taking aspirin or beta blockers are very suggestive of asthma. Moreover, symptoms that are relieved by bronchodilators point toward asthma as the underlying cause.

Symptoms that lower the probability of asthma include those of prominent dizziness, light-headedness, and peripheral tingling (all point toward dysfunctional breathing); a chronic productive cough in the absence of wheeze or breathlessness (more likely bronchiectasis); a repeatedly normal physical examination of chest when symptomatic; and voice disturbance, symptoms with colds only, or a significant smoking history. Taking a detailed patient history and recording spirometry when the patient is symptomatic is crucial to the accurate classification of the disease. There are several differential diagnoses that must be considered when making a diagnosis of asthma. These are outside the scope of this chapter. A good guide to the clinical diagnosis and management of asthma are the British Thoracic Society Asthma guidelines, which are updated regularly and are available at www.brit-thoracic.org.uk/clinical-information/asthma/.

Triggers for asthma symptoms include allergens (proteins) derived from, among others, house dust mite (HDM), cat and dog dander, cockroach, and fungi, especially *aspergillus fumigatus*. There are numerous other nonallergic symptom triggers, including weather and atmospheric change, air pollution, exercise, menstruation, emotion and laughter, viral infection, gastroesophageal reflux, and rhinitis. Various occupations can both cause and worsen asthma (e.g., welding, baking, paint spraying).

Asthma symptoms often overlap with other allergies, including oral allergy syndrome and food allergies. Delineating between causes of symptoms can be difficult. If a specific aeroallergen is identified that triggers asthma (classically tree or grass pollen), a course of either sublingual or subcutaneous desensitization can be undertaken to reduce symptoms on exposure to that specific allergen.

There are multiple other symptom triggers that can include both allergic and non-allergic mechanisms; examples include exposure to chemicals, such as perfumes, paint and bleach, plants (e.g., *Ligustrum vulgare*-privet hedges), salicylates, and sulphites.

28.2.2 Associated Comorbidities and Confounders

Multiple factors have been shown to either directly cause or worsen symptoms of asthma [4]. These are listed in Box 28.1. Other factors not directly related to the disease can worsen symptoms and are associated with more frequent exacerbations. These include dysfunctional breathing, vocal cord dysfunction, psychosocial problems, and poor adherence to treatment regimes [4].

Other forms of asthma are recognized, among them occupational, exercise-induced, and pregnancy-related asthma. These subdivisions are somewhat arbitrary, but may have different causes, prognoses, and complexities. For example, occupational asthma includes that triggered by IgE-mediated mechanisms, that due to specific agents with unclear pathophysiology, and that secondary to irritants (also known as RADS-reactive airways dysfunction syndrome). Underlying asthma can also be made worse by occupational exposure (work-aggravated).

Most experts regard exercise-induced asthma as a different disease, albeit one resulting in similar symptoms. Exercise-induced asthma is often diagnosed with different investigations, including eucapnic voluntary hyperventilation or exercise testing. It is associated with certain sports (swimming and cross country skiing in particular) and has a unique pathophysiology directly related to damage of the airway epithelium.

BOX 28.1 Factors Associated with the Triggering or Exacerbating of Asthma

Atopy/allergy (HDM, animal fur, pollens, etc.)
Viral infection
Bacterial infection
Occupation
Medication: beta blockers; nonsteroidal anti-inflammatories
ABPA
Obesity
Gastro-esophageal reflux disease
Hormonal changes: pregnancy; menstruation; menarche
Bronchiectasis
Smoking
Weather changes
Stress/emotion
Exercise

Asthma control can deteriorate in pregnancy, due to both hormonal effects and the physical effects of a gravid uterus on diaphragm and respiratory muscle function. Pregnancy can also impact treatment decisions. Leukotriene modifiers are generally contraindicated, and the health of the fetus as well as that of the mother needs to be considered.

Poor adherence is probably the biggest single issue affecting asthma control. Studies have shown that people with asthma over-report use of ICS; one study found that although 95% of responders said they used their inhalers, only 58% actually did. Adherence to ICS therapy is associated with a lower rate of death, whereas increasing use of reliever (salbutamol/bricanyl) medication, which improves symptoms but does not treat underlying airway inflammation, is associated with increased mortality [7,8].

A 2010 report entitled *Evaluation of the Scale, Causes and Costs of Waste Medicines: Report of DH funded national project* (Trueman et al. York Health Economics Consortium and University of London School of Pharmacy, 2010) put figures to the possible cost benefits from improving asthma treatment adherence in the United Kingdom. It estimated the resulting net benefit associated with compliance to be £2250 per patient, with a reduction in expected annual treatment costs of approximately £75 per patient per year. Based on an asthma point prevalence of 5.8%, the findings suggest that there are almost 1.8 million asthmatics in the United Kingdom who are noncompliant. If interventions were available to change the behavior of all partially compliant medicine users to raise the percentage to 80% or more, the report estimated that over £130 million in treatment cost savings could be realized in the United Kingdom alone. A more modest target of doubling current compliance rates would result in savings of approximately £90 million per year.

28.2.3 Asthma Control

The aim of asthma treatment is to improve symptoms, maintain lung function, and prevent exacerbations. From a patient's perspective, these are simple aims, but for the purposes of study design and treatment trials they do not address the complexity of the disease process. Accordingly, guidelines now specifically address asthma control (symptoms and lung function) and severity (need for treatment) separately. The current Global Initiative for Asthma (GINA) guidelines assess levels of control based on daytime symptoms, limitation of activities, nocturnal symptoms/awakening, need for reliever/rescue treatment, lung function (PEF or FEV_1), and exacerbations [9] (Table 28.1). Aspects of

TABLE 28.1 Global Initiative for Asthma Classification of Asthma Severity

Step	Symptoms/Day	Symptoms/Night	PEF or FEV_1	PEF Variability
Step 1 (intermittent)	Fewer than once/week; asymptomatic, normal PEF between attacks	Two or fewer/month	80% or more	Less than 20%
Step 2 (mild persistent)	More than once/week, fewer than once/day; attacks may affect activity	>2 times a month	>/=80%	20–30%
Step 3 (moderate persistent)	Daily; attacks affect activity	More than once/week	60%–80%	More than 30%
Step 4 (severe)	Continuous; limited physical activity	Frequent	60% or less	More than 30%

Note: Multiple factors, including prevalence of day/night symptoms, airflow limitation, and pulmonary function variability, categorize disease severity. PEF—peak expiratory flow; FEV_1—forced expiratory volume in one second.

control are measured by validated questionnaires, including the Asthma Quality of Life Questionnaire and the Asthma Control Questionnaire.

Treatment steps are defined by the treatment level required to maintain asthma control (Table 28.1), with patients at steps 1–2 having well-controlled asthma requiring little treatment, and patients at step 4 having poorly controlled asthma despite four or five different drugs. The step approach is similar to that used by the British Thoracic Society (BTS) Guidelines [4].

28.2.4 Exacerbations

28.2.4.1 Definition

Asthma is a variable disease both temporally and clinically. There is also overlap between symptoms and exacerbations. This overlap was the subject of a recent consensus document that defined asthma exacerbations as "events characterized by a change from the patient's previous status." Severe exacerbations are "events that require urgent action on the part of the patient and physician to prevent a serious outcome, such as hospitalization or death from asthma," and moderate asthma exacerbations are "events that are troublesome to the patient, and that prompt a need for a change in treatment, but that are not severe…clinically identified by being outside the patient's usual range of day-to-day asthma variation" [10]. The most practical definition of an asthma exacerbation is probably an episode of worsening symptoms not responding to increasing bronchodilator therapy. This definition was employed in an elegant study on the time course of peak flow changes [11]. For the purposes of clinical studies, most authors define *exacerbation* as the need for rescue courses of systemic corticosteroids (prednisolone/prednisone).

28.2.4.2 Cost

Although the majority of patients with asthma are treated and investigated in primary care settings, the main burden of asthma is due to its severe form (i.e., step 4) and exacerbations. Using the United Kingdom as an example (population 60 million), asthma is responsible for more than 1200 deaths per year and for over 50,000 hospital admissions, with an annual expenditure of £800 million on pharmaceutical costs alone.

It was recently estimated that a patient controlled at the mildest end of the spectrum (British Thoracic Society (BTS) guidelines: step 1 with no exacerbations) would cost 50 times less to provide his or her package of care than would a patient at the worst end of the spectrum (BTS guidelines: step 5 with exacerbations) [12]. It is also calculated that 5% of asthma patients are responsible for at least 50% of the total healthcare burden [4]; the most expensive individual cost is an ICU admission (BTS level 3) due to a life-threatening exacerbation. In the U.S. more than half (53%) of asthma sufferers in 2008 had an asthma attack, and of these half of the children and one-third of the adults missed school or work because of it; on average, children missed four days of school and adults missed five days of work.

28.2.4.3 Pathogenesis

Exacerbation pathogenesis is not fully understood. Although research has focused on infective agents, especially viral, there may be other explanations.

Asthma exacerbations are associated with both inflammatory and immunological cell infiltration. The inflammatory cell infiltrate is composed of varying numbers of eosinophils, neutrophils, and lymphocytes. This airway inflammation, combined with smooth-muscle hypertrophy and thickening of the lamina reticularis, is accentuated by mucus plugs, serum protein deposition, inflammatory cell and cellular debris, leading to blockage of the airway (airflow obstruction) and wheeze.

Viral Agents

Approximately 80% of exacerbations are associated with respiratory tract viral infections, with rhinovirus responsible for about two-thirds of cases [13]. Asthmatic subjects also have much more severe lower respiratory tract illness with rhinovirus infection than do healthy control subjects. The mechanism for this is not fully understood. Infection induces inflammation, increasing levels of neutrophils, eosinophils, CD41+ cells, CD81+ cells, and mast cells through increased mRNA expression and translation of IL-6, IL-8, IL-16, eotaxin, IFN-γ-induced protein 10 (IP-10), Chemokine (C-C motif) ligand 5 (CCL5/RANTES), and other proinflammatory cytokines. Other viruses implicated in exacerbations include enterovirus, coronavirus, influenza, parainfluenza, respiratory syncytial virus, metapneumovirus, adenovirus and bocavirus [14]. Although influenza vaccination is recommended for all individuals with asthma, there is currently no hard evidence that this improves outcomes.

Bacterial Agents

Bacterial infection has also been implicated in asthma exacerbation. Individuals with asthma have an increased risk of invasive pneumococcal disease [15], and an increased frequency of detection of *Chlamydophila pneumoniae* [16]. One study found *Mycoplasma pneumoniae* infection in 20% of patients hospitalized for severe asthma [17]. Several trials have evaluated the role of macrolide antibiotics in treating and preventing asthma exacerbations. One study randomized adults with exacerbations to the ketolide antibiotic telithromycin or placebo; the telithromycin group had a small but significant improvement in symptoms and lung function from exacerbation to the end of treatment [18]. Further studies of macrolide antibiotics (which have immunomodulatory as well as antibiotic properties) are under way, but as yet macrolide antibiotics are not recommended by international asthma guidelines.

New techniques for assessing the airway bacterial microbiota have established that the airways are not normally sterile, and may permit the role of bacterial infection in airways disease to be further delineated. A recent study found that pathogenic *Proteobacteria*, particularly *Haemophilus* species, were much more frequent in the bronchi of patients with asthma or COPD than in controls [19].

Allergen Exposure

Prolonged exposure to aeroallergens can result in chronic airway inflammation via Th2-driven IgE mechanisms. This immunological reaction may intensify airway inflammation, increase inflammatory cell activation, and stimulate mucus glands to hypersecrete, leading to airway obstruction. Studies of bronchoalveolar fluid before and after allergen challenge show eosinophilic inflammation as the major airway response, associated with late-phase responses of airflow obstruction. In addition, IL5 and IL13 are significantly raised. Exposure to seasonal allergens has been implicated in sudden asthma-related deaths; HDM, cat, and cockroach allergen sensitization are risk factors for emergency treatment [20]. Grass pollen sensitization, or "thunderstorm asthma," has also been associated with epidemics of exacerbations [21].

Fungal Exposure

Fungal allergens are found both outdoors and indoors, and sensitivity to them is a risk factor for the development, persistence, severity, and mortality associated with asthma. Individuals with asthma are often sensitized to fungi such as *Cladosporium* species, *Alternaria* species, *Penicillium* species, and *Candida* species [22].

Alternaria species sensitization and exposure are associated with symptoms and a 200-fold increased risk of respiratory arrest in subjects with asthma [23]. Fungal exposures may worsen disease by different mechanisms; fungal-associated proteases may lead to the development of allergic airway inflammation along with IgE-mediated responses; fungal allergen-induced asthmatic reactions evoke an IL5 response with increased eosinophil recruitment and degranulation [24].

The relationship between fungal exposure and asthma is complicated by the degree of sensitivity to fungal allergens and the resultant allergy. A specific disease entity called allergic bronchopulmonary aspergillosis (ABPA) exists in which colonization of the respiratory tract with the ubiquitous *Aspergillus fumigatus* is associated with an increased allergic response (both type 1 and type 3) and severe asthma exacerbations. Trials in both ABPA and a related disease, severe asthma with fungal sensitization (SAFS), have shown some benefit with the oral antifungal itraconazole [25], but good-quality trial data are lacking.

28.2.5 Diagnosis

Unlike other chronic conditions, there is no gold standard for the diagnosis of asthma. Although tests of airway function can be used to diagnose and study different aspects of airway disease (airway inflammation/airway hyper-responsiveness), diagnosis still depends on the presence of specific symptoms, which include wheeze, cough, chest tightness, and difficulty breathing. These symptoms are often variable, worse at night, or worse in response to triggers. They often respond to asthma therapy and may be associated with atopy/allergy and a family history of similar problems.

The lack of a gold standard affects clinical provision and studies alike; most estimates suggest that asthma is wrongly diagnosed in approximately 30% of patients [26]. The development of a cheap and reliable noninvasive test that can diagnose asthma is seen as the holy grail for diagnosis-based research.

28.2.5.1 *Airflow Limitation*

Measurements of airflow limitation provide an assessment of the severity, reversibility, and variability of airflow obstruction, and can help confirm the diagnosis of asthma. Spirometry (Forced Expiratory Volume in 1 second (FEV_1), Forced Vital Capacity (FVC), and FEV_1/FVC) is the preferred method of measuring airflow limitation, as it is more repeatable and less effort dependent than peak flow. A FEV_1/FVC ratio of less than 0.7 is consistent with a diagnosis of airflow obstruction, and an increase in FEV_1 of >12% (or >200 ml) after administration of a short-acting bronchodilator indicates reversible airflow limitation consistent with asthma [27]. Importantly, most asthma patients will not exhibit reversibility at each assessment, and repeated testing is advised to confirm a diagnosis. Consequently, an absent response to bronchodilators does not exclude asthma.

Peak expiratory flow (PEF) measurements can be used to monitor asthma. They are ideally compared to the patient's own previous best measurements using his or her own peak flow meter. An improvement of 60 L/min (or >20% of the prebronchodilator PEF) after inhalation of a bronchodilator, or diurnal variation in PEF of more than 20% (with twice-daily readings, more than 10%), suggests poorly controlled asthma.

28.2.5.2 *Other Diagnostic Approaches*

Other approaches have been utilized to diagnose asthma, both to confirm a formal diagnosis and to identify corticosteroid response. These approaches include noninvasive measurements of airway inflammation or assessment of airway hyper-responsiveness. Measuring airway inflammation is relatively easy, either using techniques to induce sputum and then counting the differential eosinophil and neutrophil cell numbers present, or by measuring the fraction of nitric oxide (NO) present in the exhaled breath (F_ENO) using portable chemical analyzers.

Airway hyper-responsiveness can be assessed by several different methods. All are designed to provoke bronchoconstriction, which is measured using spirometry. The concentration (or dose) at which bronchoconstriction occurs is then used to categorize the degree of airway responsiveness and the presence or absence of asthma. Although these tests are widely employed in research settings, their use in a clinical setting is hampered by their varying sensitivity and specificities for asthma diagnosis [4], as both airway inflammation and airway hyper-responsiveness can be present in healthy individuals without symptoms.

28.2.5.3 *Pediatric Asthma*

The symptoms of asthma in children are recurrent wheezing, cough, difficulty breathing, and chest tightness. Evaluation of these symptoms is critical to the diagnosis, treatment, and outcome measures used in clinical studies; however, asthma triggers and inflammatory cell type are increasingly used to define childhood asthma phenotypes.

Pediatric asthma is complicated by wide variations in symptom prevalence. The majority of children with asthma have mild or moderate disease; 5% of all asthmatic children have chronic symptoms and/or recurrent exacerbations, despite maximum treatment with conventional medications [28]. Asthma in childhood is heterogeneous in several ways, including etiology, clinical presentation, and response to treatment. Several attempts have been made to stratify treatment on the basis of different asthma phenotypes in children, with varying degrees of success. Stratifying treatments in pediatric asthma is more complicated because of patient age and the inability (dependent on age) to perform more complicated or invasive tests.

Etiology

In children, the onset and progression of asthma result from a complex interplay between genetic background and environmental exposures. The complexity is confounded by wheezing illnesses consisting of several distinct disease entities, with no general agreement on their number or underlying mechanisms [29].

Among the environmental factors that influence the risk of asthma are viral and bacterial respiratory infections. Exposure to environmental tobacco smoke is also associated with increased rates of early viral illnesses. Other factors associated with the risk of wheeze include physical factors associated with increased breathing (exercise, laughing, crying, and excitement) or allergens (aeroallergens and food allergens). There are considerable age-related changes in the relative importance of trigger factors for wheeze in children.

Human rhinovirus (HRV) has been implicated in both the etiology of asthma and exacerbations; infants who have HRV infections with wheezing are at a significantly increased risk for subsequent asthma, and over 50% of childhood asthma exacerbations are triggered by HRV. It is important to note that exposure to HRV does not lead to wheezing illness in all children, nor does wheezing illness result in asthma in all cases, suggesting that the host genotype also plays a role. A recent study of two cohorts of children with asthma found that variants at the 17q21 locus were associated with asthma in children who had HRV wheezing illnesses, implicating an interaction between HRV wheezing illness in early childhood and the 17q21 genotype (see Section 28.3.2).

Clinical Presentation

The presentation of asthma symptoms and exacerbations in childhood also varies. Some children have frequent exacerbations with few daily symptoms, while others have recurrent symptoms without airway inflammation and exacerbations. Other diseases coexist with childhood asthma, including gastroesophageal reflux disease, severe asthma with fungal sensitization, obesity, and vocal cord dysfunction. Targeting these coexisting disorders has met with varying clinical success [30].

28.3 ASTHMA PATHOPHYSIOLOGY

28.3.1 Biological Basis

An important component of the development of asthma is the inflammatory cascade. This involves the infiltration of a number of inflammatory cells, such as eosinophils, neutrophils, B- and T-lymphocytes, macrophages, mast cells, dendritic cells, and basophils, into the airway, and the release of inflammatory mediators (e.g., leukotrienes, histamine, cytokines, and chemokines) by airway, structural cells (including epithelial, smooth-muscle, endothelial, and fibroblast cells). However, since the degree of inflammation in asthma is not directly related to asthma severity, this suggests that other factors, such as structural changes in the airways, also play a role in disease modulation and progression. These structural changes have been termed *airway remodeling* and may exist in the presence or absence of inflammatory mechanisms in the airway [31].

28.3.1.1 Inflammatory Response

The inflammatory response in asthma is a result of excessive activation of mast cells in the airway, where in the early response to allergen challenge, degranulation of mast cells results in the release of a number of proinflammatory factors such as IL4, leukotrienes, and histamine. This causes an immediate hypersensitivity response that in turn leads to airway narrowing. The concomitant release of other inflammatory factors such as cytokines and chemokines from the same mast cells provides an optimal environment for recruitment to the airways of other inflammatory cells such as eosinophils, basophils, neutrophils, and T-lymphocytes.

Once in the airways, these cells act in tandem with allergen-activated macrophages, resulting in what is termed the late asthmatic response (occurring at 4–8 hours post-challenge). Here, through the action of a number of released potent immunomodulators (e.g., TNFα, IL3, IL4, IL5, and IL13), airway narrowing occurs. This may occur in periods that last more than 24 hours. Of the inflammatory cells involved in the asthmatic response, it is the infiltration of eosinophils that is considered characteristic of the asthmatic phenotype, particularly that of mild/moderate allergic asthma (Figure 28.1). For a more extensive description of inflammatory mechanisms underlying allergic asthma see Hodge & Sayers [32].

Although inflammation in asthma has been extensively studied, the relationship between inflammation and clinical symptoms of asthma is still unclear. This may be partially explained by a possible link between airway inflammation and airway hyper-responsiveness [33]. Here, airways become more susceptible to sensitizing agents, which would otherwise be unable to trigger an airway response, because of (a) increased release of mediators such as histamine and leukotrienes, (b) abnormal smooth-muscle behavior, and (c) airway thickening.

28.3.1.2 Airway Remodeling

Airway remodeling plays an important role in asthma pathogenesis. Airway structural changes characteristic of airway remodeling occur because of prolonged airway inflammation. Here, prolonged release of inflammatory factors results in:

- Thickening of the smooth-muscle bronchial walls, leading to airway narrowing
- Denudation of the bronchial epithelium
- Hyperplasia and hypertrophy of the airway smooth-muscle
- Hyperplasia and hypertrophy of the epithelial goblet cells, resulting in the formation of large mucus plugs liable to occlude the airway and increase airway hyper-responsiveness

Airway remodeling also involves changes occurring in the extracellular matrix and its constituent proteins (collagens I/III/IV, fibronectin, and laminin) and structural changes such as angiogenesis, vasodilation, increased airway blood flow, and changes in autonomic neurological function [34].

The development of asthma results in a degree of airway remodeling. The structural changes that define airway remodeling are an important feature in asthmatic airways, where even in newly diagnosed asthma patients, a degree of structural change can be identified in the bronchial wall [35]. Airway remodeling can be defined as a process of sustained disruption and modification of structural cells and tissues leading to the development of a new airway wall morphology [36]. Airway remodeling (Figure 28.2) is initiated either through various inflammatory pathways, highlighting the importance of the inflammatory pathway [37], or through bronchial hyper-responsiveness, where airway remodeling occurs independently of inflammation [38].

28.3.2 Genetic Basis

Our understanding of the extent and nature of genetic variation in the human genome has dramatically improved since the publication of the first draft genome sequence in 2001,

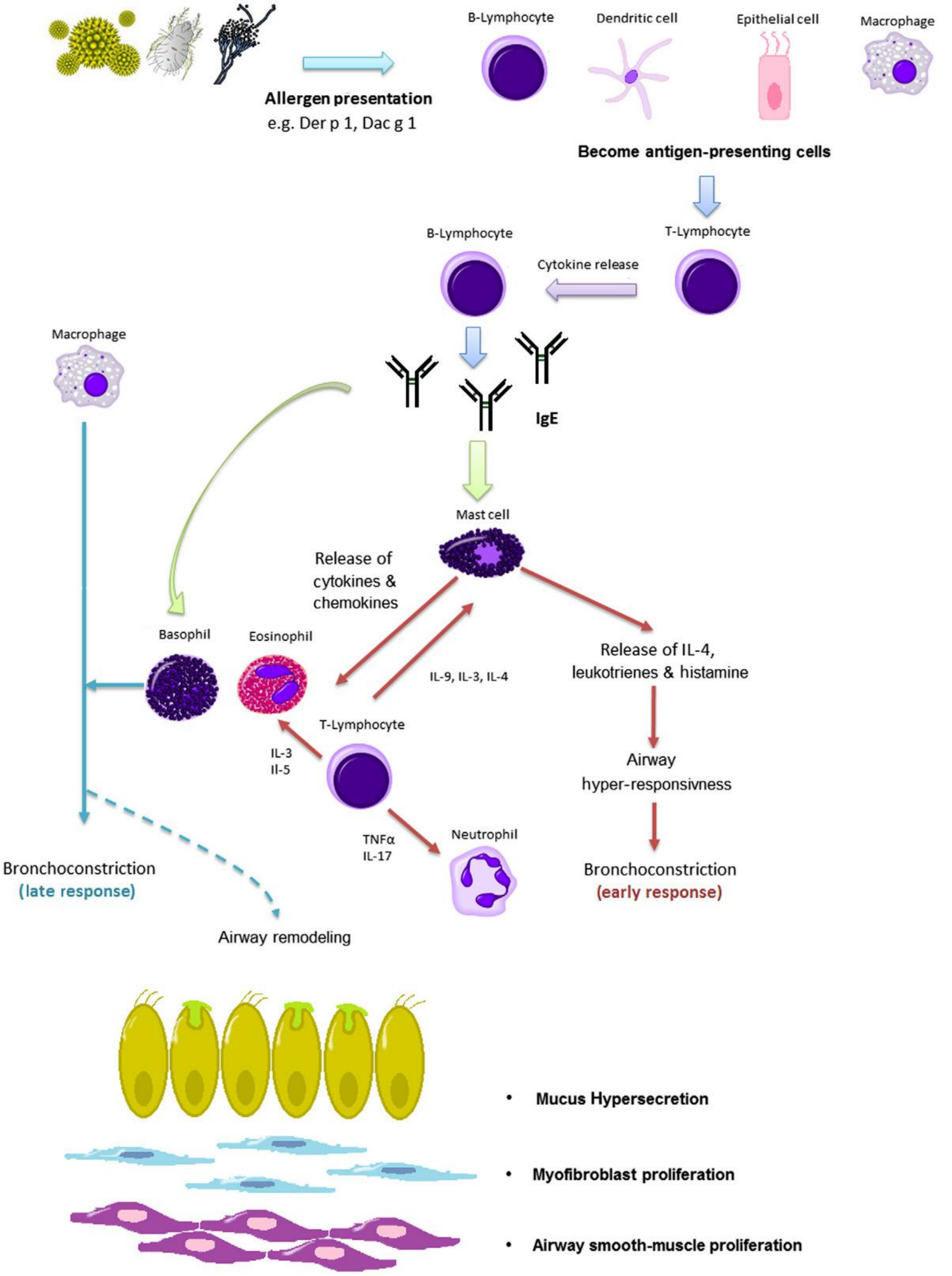

B-Lymphocyte
Dendritic cell
Epithelial cell
Macrophage
Allergen presentation
e.g. Der p 1, Dac g 1
Become antigen-presenting cells
B-Lymphocyte
T-Lymphocyte
Cytokine release
Macrophage
IgE
Mast cell
Release of cytokines & chemokines
Basophil
Eosinophil
IL-9, IL-3, IL-4
Release of IL-4, leukotrienes & histamine
T-Lymphocyte
IL-3
Il-5
Airway hyper-responsivness
TNFα
IL-17
Neutrophil
Bronchoconstriction
(late response)
Bronchoconstriction
(early response)
Airway remodeling
Mucus Hypersecretion
Myofibroblast proliferation
Airway smooth-muscle proliferation

FIGURE 28.1 **Asthma pathophysiology.** Exposure to allergens—for example, *Dermatophagoides pteronyssinus* 1 (Der p1) and *Dactylis glomerata* (Dacg1)—originating from various sources, including pollen, house dust mite, and mold, are taken up by dendritic cells, B-lymphocytes, epithelial cells, and macrophages, which present the antigens to T-lympocytes. This activates the T-lymphocytes to produce cytokines that regulate immunoglobulin E (IgE) production by the B-lymphocytes. Bound IgE activates mast cells in the airways. Activated mast cells initially release a number of factors, including IL4, leukotrienes, and histamine, that cause airway hyper-responsiveness and result in an early response to allergen stimulus via bronchoconstriction. The concurrent release of cytokines and chemokines from the mast cell recruits eosinophils, basophils, and T-lymphocytes to the airway. These cells, in association with T cell-activated neutrophils and with antigen stimulated macrophages, release chemokines and leukotrienes over an extended period of time, resulting in a late response to allergen stimulus via bronchoconstriction. Prolonged stimulation of the late-response factors ultimately leads to the airway structural changes such as mucus hypersecretion, myofibroblast, and airway smooth-muscle proliferation, which are common to the asthmatic lung.

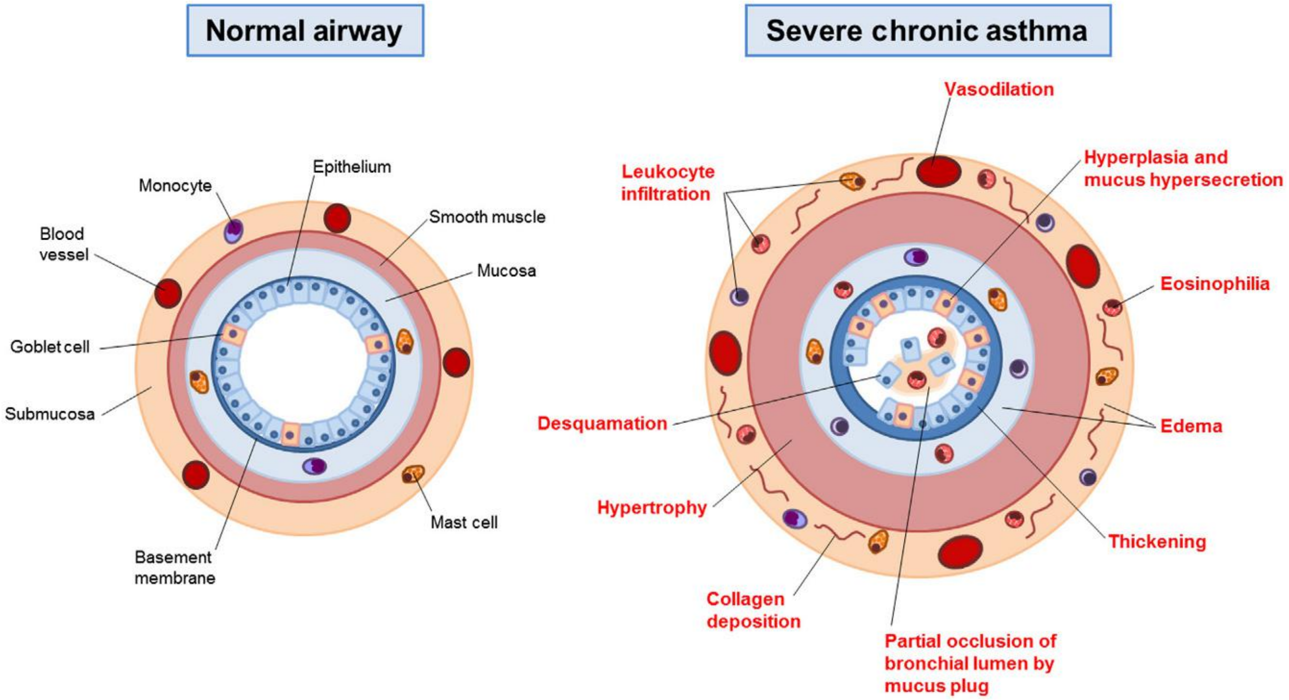


FIGURE 28.2 **Asthmatic airway pathology.** This schematic comparison of a normal airway with that observed in severe chronic asthma indicates histological changes that accompany recurring inflammation seen over time. Unlike that from the unaffected individual, the bronchial mucosa from the severe asthmatic displays thickening of the basement membrane, airway smooth-muscle hypertrophy, leukocyte infiltration, epithelial cell desquamation, goblet cell hyperplasia in the epithelial lining accompanied by mucus hypersecretion, and plugging of the bronchial lumen, as well as edema and collagen deposition in the submucosal area. (Not drawn to scale.) *Source: Reproduced with permission from Hodge and Sayers* [32].

and it continues to improve exponentially with initiatives such as the 100,000 Genomes project, which aims to genome-sequence this number of UK subjects, and the Encyclopedia of DNA Elements Consortium (ENCODE), an initiative to identify the genome's regulatory regions. Recent figures suggest that more than 6.9 million single-nucleotide polymorphisms (SNPs) or single-base-pair changes exist in a genome consisting of 3 billion base pairs. Similarly there is a growing realization that deletions, insertions, and expansions of tandem repeats also represent significant variation. These genetic polymorphisms are found throughout the genome and have the potential to influence gene function in several ways leading to human disease: (1) a coding-region nonsynonymous polymorphism can alter the amino acid sequence and structure/function of the protein; (2) a polymorphism can introduce a stop codon leading to the production of a nonfunctional protein; and (3) a polymorphism in a regulatory region can regulate the expression levels of a gene.

28.3.2.1 A Complex Genetic Disorder

It has long been known that atopic diseases such as asthma run in families. In 1916, using 621 atopic probands and 76 nonatopic controls and their respective families, it was shown that 48.4% of atopic probands had a family history of atopy, compared with 14.5% in the control population [39]. Similarly, a study of 176 families showed a very high concordance of asthma, hayfever, and eczema in parents and children [40]. Twin studies have been useful in identifying a significant concordance of asthma that is higher in monozygotic twins (identical genotype) versus dizygotic twins (on average sharing half of their genes). In the largest study to date, using 11,688 Danish twin pairs, it was suggested that 73% of susceptibility to asthma was genetic, with a substantial environmental component [41]. More recent studies suggest hereditability estimates of 35–95% for asthma [42].

Therefore, asthma is considered a complex genetic disorder and, in contrast to single-gene disorders (e.g., cystic fibrosis), involves multiple genes, with expression influenced by both genetic and environmental factors. Multiple environmental factors are important in asthma development, including tobacco smoke exposure, respiratory viral infections, antibiotic use, diet, and allergen exposure. Gender and ethnic background also make a significant contribution. This complex mode of inheritance combined with the heterogeneity in the presentation of the disease has made gene discovery a challenge. Early studies investigated inheritance through families containing multiple affected children, using linkage analyses and candidate gene approaches based on biology or location in the genome. However, the

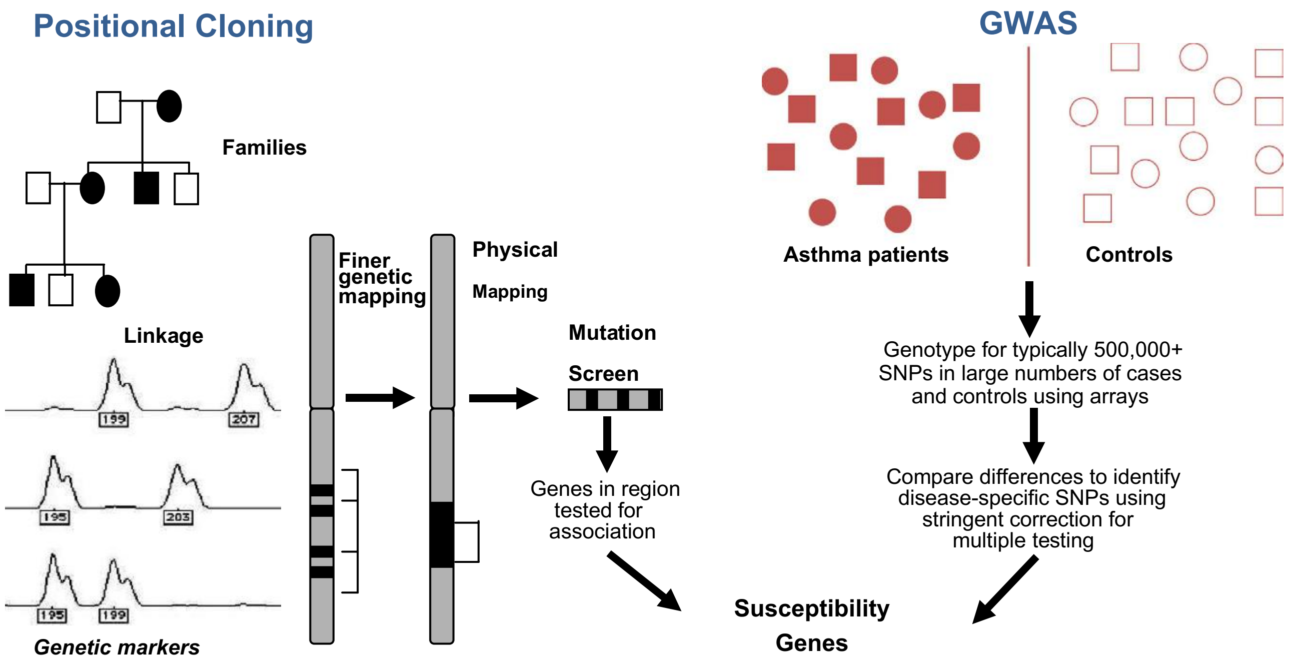

Positional Cloning
Families
Linkage
199
207
195
203
195
199
Genetic markers
Finer genetic mapping
Physical Mapping
Mutation Screen
Genes in region tested for association
GWAS
Asthma patients
Controls
Genotype for typically 500,000+ SNPs in large numbers of cases and controls using arrays
Compare differences to identify disease-specific SNPs using stringent correction for multiple testing
Susceptibility Genes

FIGURE 28.3 **Asthma gene discovery methods.** Positional cloning involves linkage analyses, which follow the transmission of genetic information through families with multiple affected children, and is followed by fine association mapping. GWAS look at the frequency of a large number of common variants between cases and controls. Both approaches lead to novel gene discovery. (Candidate gene approaches omitted for clarity.)

reproducibility of these findings was disappointing primarily because of inadequate power, subject heterogeneity (different phenotype definition), population stratification, and multiple testing without correction.

More recently, hypothesis-free genome-wide association studies (GWAS), which examine association with typically 500,000+ common (>5% frequency) polymorphisms spanning the genome in cases and controls, using very stringent statistical thresholds (Figure 28.3), have been very successful. Via these multiple approaches, over 190 asthma genes have been described in more than 1000 publications. The major reproducible findings for asthma follow below (for more comprehensive reviews see [42,43]). Although this chapter focuses on the genetics of asthma diagnosis, subphenotypes in asthma (e.g., atopy, serum total IgE levels, and lung function (e.g., FEV_1)) have also been investigated for genetic influences.

28.3.2.2 Genes Identified by Positional Cloning (1980s and 1990s)

Positional cloning involves linkage analyses (region or whole genome) followed by fine mapping using association. Linkage analysis uses family data to follow the transmission of genetic information spanning the entire genome (commonly short tandem repeat markers) across generations. These data are used to determine if a genetic marker is close to, or linked with, a gene involved in a particular disease using families with multiple affected children. Once a specific chromosomal region has been identified in this hypothesis-free approach, SNPs spanning the region are tested to see if they are more common in people with asthma compared to people without. The first gene to be identified with confidence using positional cloning was disintegrin and metalloprotease 33 (*ADAM33*) [44].

Using a cohort of 460 families, linkage was demonstrated to a locus on chromosome 20p13 for asthma, bronchial hyper-responsiveness (BHR), and total IgE levels. Subsequently, polymorphisms spanning *ADAM33* were associated with asthma in the second stage. *ADAM33* is thought to contribute to airway remodeling via its enzymatic functions. Using positional cloning, multiple genes have been identified with functions varying from transcription factors to epithelial differentiation and tissue remodeling. These genes have provided a novel insight into potential mechanisms that are altered in asthma (Table 28.2).

Polymorphisms spanning the urokinase plasminogen activator receptor (*PLAUR* or *UPAR*) were also associated with asthma susceptibility and were associated with rate of decline of lung function in asthma through linkage/association analyses on chromosome 19q13 [45]. uPAR is a serine protease receptor involved in multiple mechanisms, including cell proliferation, migration, and extracellular matrix degradations via plasmin generation. These data suggest that genetic factors may be important in determining airway structural changes or “remodeling” in asthma [46].

TABLE 28.2 Asthma Susceptibility Genes Identified Using Positional Cloning, Candidate Gene, and Genome-Wide Association Studies

Gene	Chromosome	Potential Function
Positional cloning		
DPP10	2q14	Involved in cleavage of dipeptides from cytokines
PCDH1	5q31.3	BHR, cell–cell adhesion, epithelial cell differentiation
HLAG	6p21	Regulates responses to inhaled allergens
NPSR1	7p15	Role in ECM expression
PHF11	13q14	Regulator of transcription
PLAUR	19q13	Epithelial function, tissue remodeling
ADAM33	20p13	Potential role in BHR/airway remodeling
Candidate gene		
IL10	1q31	Regulatory T cell function, anti-inflammatory cytokine
CD14	5q31.1	Innate immune responses
IL4	5q31.1	IgE-isotype switch, T cell differentiation
IL13	5q31	Inflammatory cytokine
ADRB2	5q32–34	Muscle relaxation, β_2 adrenergic receptor agonist efficiency
LTA	6p21.3	Immune modulation
HLA-DRB1	6p21.3	Regulate responses to inhaled allergens
HLA-DQB1	6p21.3	Regulate responses to inhaled allergens
TNFA	6p21.3	Inflammatory cytokine
FCER1B	11q12.1	IgE-mediated responses, mast cell degranulation
GSTP1	11q13	Oxidative stress
IL18	11q22.2	Inflammatory cytokine
STAT6	12q13	Transcription factor in IL4 receptor signalling pathway
IL4RA	16p12	IL4 receptor pathway
CCL5	17q11–12	Chemokine, inflammatory cell recruitment, ASM proliferation
TGFB1	19q13.1	Pro-fibrotic cytokine
GWAS		
IL6R	1q21	Regulatory T cell function, T cell differentiation
DENND1B	1q31	Memory T cell functions
IL1RL1	2q11	IL33 receptor, recruitment/activation of inflammatory cells
PDE4D	5q12	Modulates cAMP, role in cell signaling, ASM function

(*Continued*)

TABLE 28.2 Continued

Gene	Chromosome	Potential Function
TSLP	5q22	Activates dendritic cells, role in Th2 immune responses
HLA-DQ	6p21	Cell membrane alloantigens, T cells responses
IL33	9p24	Recruitment/activation of inflammatory cells
LRRC32	11q13	Regulatory T cell function
SMAD3	15q22	TGFβ signaling intermediate, fibrosis
*ORMDL3/GSDMB**	17q21*	Sphingolipid synthesis (ORMDL3), cell apoptosis? (GSDMB)
IL2RB	22q12	Binds IL2/IL15, lymphoid cell differentiation

BHR: bronchial hyper-responsiveness; ECM: extracellular matrix; ASM: airway smooth muscle; IgE: immunoglobulin E; IL: interleukin.

**The 17q21 locus was specifically associated with childhood onset asthma in GWAS. See main text for abbreviations and key references.*

Note: For candidate gene approaches, genes focused to replication in more than 10 studies [47,48]. For GWAS, genes focused to those meeting conventional genome-wide significance ($p < 5 \times 10{-}8$) and/or independent replication in the Caucasian population.

28.3.2.3 Genes Identified in Candidate Gene Studies (1970s to the Present)

Candidate gene studies are hypothesis-driven and based on the suggested biology of a gene product or the location of the gene in a chromosomal region previously linked to asthma. Such studies commonly employ asthma cases versus controls, although family-based association has also been used. For excellent reviews of candidate gene studies in asthma, see Ober and Hoffjan [47] and Undarmaa et al. [48]. The primary highly reproducible genes identified using candidate gene approaches are multiple components of the interleukin 4/13 axis (see Table 28.2), including the cytokine genes themselves (*IL4/IL13*); related receptor (*IL4RA*) and downstream signaling effector signal transducers and activators of transcription (*STAT*) 6. The role of the *IL4/IL13* axis in the pathogenesis of asthma has been extensively documented [49].

Human studies have identified elevated numbers of cells expressing IL13 mRNA in the bronchial tissue of atopic and nonatopic asthmatic subjects [50]; administration of recombinant IL13 in mouse lungs resulted in an increase in airway mucus secretion, development of subepithelial fibrosis, airway hyper-responsiveness (AHR), and eosinophilic airway inflammation—that is, several key features of the human disease [51]. IL13 is produced by a variety of cells, including Th2 CD4+, Th1 CD4+, CD8+ T cells, mast cells, basophils, and eosinophils. IL13 mediates its effects by interacting with a complex receptor system comprising an IL-4Rα/IL13Rα1 heterodimer and the IL13Rα2 receptor [49]. The finding that genetic polymorphisms within *IL13* that determine levels and/or structure and are associated with the development of asthma make biological sense.

28.3.2.4 GWAS in Asthma (2007 to the Present)

The use of GWAS approaches to identify asthma susceptibility genes has revolutionized our understanding of asthma genetics. While positional cloning showed great success with simple Mendelian disorders, it did not perform well for complex diseases. Similarly, candidate gene studies are useful but very inefficient. The capacity to interrogate the entire genome for 500,000+ common SNPs in cases and controls has provided an excellent platform for gene discovery.

The first GWAS for asthma used a discovery cohort of 994 patients with childhood onset asthma and 1243 non-asthma controls, and identified significant association to a locus on chromosome 17q21 [52]. This locus includes genes for *zona pellucida* binding protein 2 (*ZPBP2*), gasdermin B (*GSDMB*) and orm1-like protein 3 (*ORMDL3*). The 17q21 locus has been reproduced as a locus associated with childhood onset asthma in many studies; however, the identification of the specific gene(s) underlying these effects remains to be resolved. *ZPBP2*, *GSDMB*, and *ORMDL3* have been implicated in gene transcription, cell apoptosis, and sphingolipid synthesis, respectively.

This initial GWAS has now been superseded by several larger-scale studies. In particular, the GABRIEL consortium study, involving 10,365 asthma cases and 16,110 controls, identified association between polymorphisms spanning *IL33*, *IL1RL1/IL18R1*, *HLA-DQ*, *SMAD3*, and *IL2RB* [53].

To date, two of the most reproducible association signals are to the IL33 gene on chromosome 9p24 and the IL33 receptor (*IL1RL1* or *ST2*) gene on chromosome 2, both of which highlight the significance of this ligand/receptor system in asthma. Interestingly, these loci are associated with severe asthma as well [54]. There is good evidence from biology that the *IL33/ST2* axis may be of relevance in asthma; mice induced to develop allergic airway disease were treated with an antibody that blocks IL33 binding to its receptor. This treatment was shown to block many of the features of the disease, including reduced serum IgE and airway eosinophil and lymphocyte counts. Significantly, both induction and resolution of the disease were attenuated [55].

In human studies, IL33 has been shown to be elevated in the airways of asthma patients, particularly in the airway structural cells, including the bronchial epithelium, and is induced in these cells by relevant stimulations (e.g., HDM [56]). In addition, a soluble form of the ST2 receptor has been shown to be elevated during asthma exacerbation [57]. Overall, it is thought that the IL33/ST2 axis may be particularly important in attracting and activating inflammatory cells in the airways; however, the precise role(s) of this pathway remains to be resolved.

As outlined in Table 28.2, other genes have been identified using GWAS, including those involved in diverse roles such as inflammatory cell function and airway smooth-muscle contraction, again providing a unique insight into potentially altered mechanisms in asthma. However, most GWAS determined single variants only confer a very modest odds ratio of 1.1–1.5 of developing asthma. For a more comprehensive review of GWAS findings in asthma, see Akhabir and Sandford [43].

28.3.2.5 Reproducibility and Missing Hereditability

Over the last 40 years, there has been great progress in identifying asthma susceptibility genes. Current and future approaches include meta-analyses using cohorts of tens of thousands of asthma patients, the incorporation of gene expression/SNP analyses (the expression quantitative trait loci (eQTL)), the investigation of asthma subphenotypes using GWAS, and whole-genome sequencing.

The real challenge is the translation of these genetic findings into a deeper understanding of the biology of asthma and the potential identification of therapeutic targets. Already several of these newly identified genes (e.g., *IL33* and IL33 receptor (*IL1RL1/ST2*)) represent excellent pharmacological targets, and it is anticipated that genetics in these loci will be essential to identifying patients most likely to gain clinical benefit from targeting this pathway.

The lack of concordance between approaches (e.g., linkage versus GWAS) can be explained by the fact that the methodologies are designed to detect different types of variants, e.g., linkage analysis has good power to detect high-risk disease-causing alleles but is not effective at identifying common alleles of modest effect, as GWAS does. It is reassuring that many of the genes identified in candidate gene approaches have been reproduced in GWAS (e.g., the *IL13/IL4* locus on chromosome 5q31, a region that has also been linked to asthma). While there has been much success in recent years in identifying genes for asthma using GWAS, the overall genetic variation accounted for by these genetic polymorphisms is very small, leading to the concept of "missing hereditability." Possible explanations for missing hereditability include: (1) rare variants (<5% frequency) with larger effect size not measured on existing platforms; (2) structural variation (e.g., copy number variation); (3) gene–environment contributions; (4) gene–gene interactions; (5) epigenetic mechanisms; and (6) overestimation of initial hereditability.

28.4 CURRENT MANAGEMENT OF ASTHMA

The aim of treatment is to achieve control of asthma and prevent exacerbations. Most guidelines adopt a stepwise approach to the management of asthma and advise stepping up treatment as necessary and stepping down when control is achieved for at least three months. Although there are several pharmacological options for asthma treatment (see Figure 28.4), the two main classes remain bronchodilators (short- or long-acting) and corticosteroids.

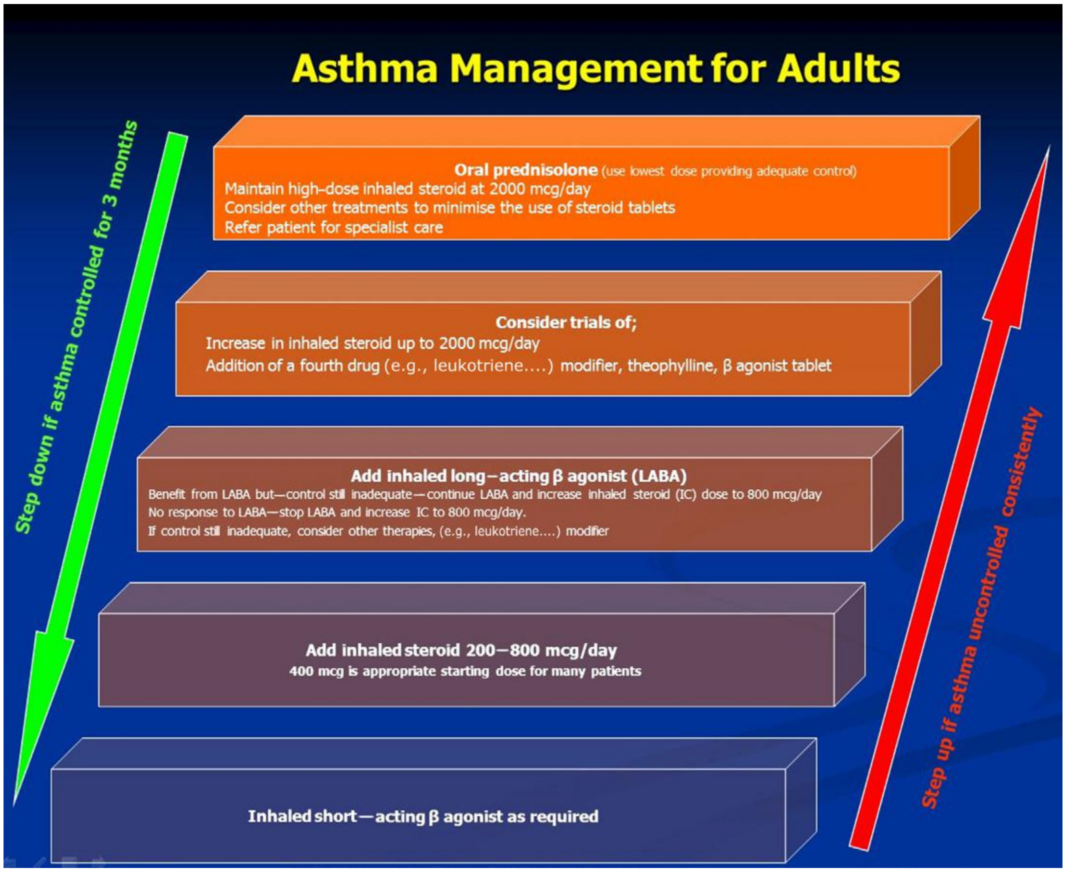
28.4.1 Bronchodilators

Short-acting bronchodilators (e.g., salbutamol/terbutaline) are sympathomimetics (β_2-adrenergic receptor agonists) that relax airway smooth-muscle, enhance mucociliary clearance, decrease vascular permeability, and possibly modulate mediator release from mast cells. They are used for symptom control, acute exacerbations, and exercise-induced asthma. Frequent or regularly scheduled use of rapid-acting inhaled β_2 agonists for long-term management of asthma does not adequately control symptoms, PEF variability, or airway inflammation.

Long-acting bronchodilators (e.g., salmeterol/formoterol) activity is identical to that of short-acting β_2-adrenergic receptor agonists, but should be used in combination with inhaled corticosteroids because of concerns regarding increased mortality when used as a monotherapy. Long-acting bronchodilators are normally introduced at treatment step 3.

28.4.2 Inhaled Corticosteroids

Inhaled corticosteroids (ICS) form the backbone of treatment for all but the very mildest asthma. ICS improve lung



Asthma Management for Adults

Oral prednisolone (use lowest dose providing adequate control)
Maintain high-dose inhaled steroid at 2000 mcg/day
Consider other treatments to minimise the use of steroid tablets
Refer patient for specialist care

Consider trials of;
Increase in inhaled steroid up to 2000 mcg/day
Addition of a fourth drug (e.g., leukotriene….) modifier, theophylline, β agonist tablet

Add inhaled long–acting β agonist (LABA)
Benefit from LABA but—control still inadequate—continue LABA and increase inhaled steroid (IC) dose to 800 mcg/day
No response to LABA—stop LABA and increase IC to 800 mcg/day.
If control still inadequate, consider other therapies, (e.g., leukotriene….) modifier

Add inhaled steroid 200–800 mcg/day
400 mcg is appropriate starting dose for many patients

Inhaled short–acting β agonist as required

Step down if asthma controlled for 3 months

Step up if asthma uncontrolled consistently

FIGURE 28.4 The current stepwise approach to asthma management.

function, symptoms, and quality of life, reduce exacerbations, and improve mortality. They have broad anti-inflammatory and immunosuppressive effects. Corticosteroids enter the cell cytoplasm and bind with the inactive glucocorticoid receptor complex. The activated glucocorticoid receptor binds to DNA at the glucocorticoid response element sequence and promotes synthesis of anti-inflammatory proteins (transactivation), and inhibits transcription and synthesis of many proinflammatory cytokines (transrepression). They also reduce the number of T-lymphocytes, dendritic cells, eosinophils, and mast cells in the airway, and reduce inducible nitric oxide production. ICS (beclomethasone dipropionate/fluticasone propionate/budesonide/flunisolide/ciclesonide/mometasone) all have side effects related to their transactivation properties, including local effects on the oropharynx (hoarseness, candidiasis, cough, and dysphonia) and systemic effects (Cushing syndrome, osteoporosis, cataracts, dermal thinning and bruising, adrenal insufficiency, and growth suppression) in children [58]. Consequently, the lowest dose required to control symptoms and prevent exacerbations should be prescribed.

28.4.3 Other Pharmacological Treatments and Alternative Therapies

Other medications are normally reserved for second- or third-line treatment. These include methylxanthines (theophylline), leukotriene modifiers (montelukast), inhaled anticholinergics (ipratropium bromide/tiotropium), and sodium chromoglycate.

A wide range of interventions have been assessed for both the primary prevention of asthma and the secondary prophylaxis of the disease once present. These treatments have varying effects, but may be recommended in individual cases. Primary prevention measures include aeroallergen and food allergen avoidance, fish oil supplementation, avoidance of tobacco smoke and air pollutants, immunotherapy, and immunization.

Measures that have been assessed for secondary prophylaxis include HDM and allergen avoidance, smoking reduction, subcutaneous and sublingual immunotherapy, dietary manipulation, weight loss, breathing techniques, and exercise training.

28.4.4 Specific Treatment for Severe Asthma

Some investigators argue that severe asthma is an inflammatory condition separate from mild/moderate asthma, rather than simply the severe end of the disease spectrum. This is difficult to prove, as underlying asthma may be complicated by other aspects (obesity/reflux/smoking/poor adherence), making it more difficult to treat and control. There are different definitions of severe asthma, but all require high-dose inhaled corticosteroids or oral corticosteroids and include recurrent exacerbations and poor control.

As high-dose ICS and oral corticosteroids are associated with side effects attempts have been made to design treatments to either reduce the amount of corticosteroid prescribed (steroid-sparing agents) or treat the underlying asthma in a different manner. Steroid-sparing agents trialled include methotrexate, gold, cyclosporin, and azathioprine. Currently, none have been recommended in any asthma guidelines. Box 28.2 lists other treatments used on an individual basis.

BOX 28.2 New and Emerging Therapies Reserved for Severe Asthma

Anti-IgE
Bronchial thermoplasty
Macrolide antibiotics
Antifungals (Itraconazole/Voriconazole)
Anti-IL5
Anti-IL13

28.5 STRATIFIED MEDICINE FOR ASTHMA

The need for a stratified-medicine approach to asthma has become apparent over the last decade. This has followed the realization that the term *asthma* covers a whole range of disease phenotypes. As techniques for investigating and researching asthma have evolved, it has become clear that different disease processes are present that require different treatments and approaches. Whether this represents disease phenotyping or just a better understanding of disease pathogenesis is unclear. Over the years, research focus has moved from the airway smooth-muscle via airway inflammation to airway immunology, and on to the epithelial–mesenchymal trophic unit. Future insight into disease pathogenesis and new treatment developments will follow with the use of the latest technologies, which include “omics” platforms and assessment of airway microbiota.

Improving our understanding of stratified medicine for asthma—targeting the right therapy to the right patient—aims to reduce hospital admissions due to poorly controlled asthma and fatalities. Although most asthma can be controlled by ICS and bronchodilators, there remains a cohort of patients with persistent symptoms, recurrent exacerbations, and worse quality of life. The current stepwise management approach (see Figure 28.4) is inefficient, particularly in light of heterogeneity in clinical presentation and response to existing therapies, making a stratified approach essential.

28.5.1 Heterogeneity in Asthma

28.5.1.1 Adult Asthma

Early work conceptualized asthma as extrinsic (allergic) and intrinsic (nonallergic) [59]. However, the recent use of unbiased approaches to classify disease using three large datasets and cluster analysis [60–62] has highlighted the different types of disease under the umbrella term asthma (Figure 28.5). Although these early studies were performed in different parts of the world and had statistical variations, their results were similar. Age at disease onset was found to be a key differentiating factor; early-onset disease was associated with more atopic and allergic conditions, whereas later-onset disease was associated with eosinophilic inflammation and obesity. These studies were all limited by a cross-sectional design and by the fact that treatment may have determined clusters. Longitudinal studies initiated in childhood are ongoing. For example, PACMAN (formally, “Pharmacogenetics of Asthma Medication in Children: Medication and Anti-Inflammatory Effects”) aims to differentiate children with uncontrolled asthma, despite ICS treatment, using a range of clinical and cellular markers [63].

28.5.1.2 Allergy

The role of allergy (an inappropriate and harmful immune response to a normally nonharmful substance which requires sensitization) and that of atopy (the tendency to develop IgE antibodies to commonly encountered environmental allergens by natural exposure in which the route of entry is across intact mucosal surfaces) have been well studied in asthma. People with extrinsic asthma were thought to develop the disease earlier in life, be atopic, and have identifiable allergic triggers as well as other allergic diseases such as rhinitis or eczema. Intrinsic asthma was thought to develop later in life (after 40 years of age), and be associated with aspirin-exacerbated respiratory disease but not with allergic sensitization. When small studies in humans suggested that levels of Th2 cytokines were similar in extrinsic

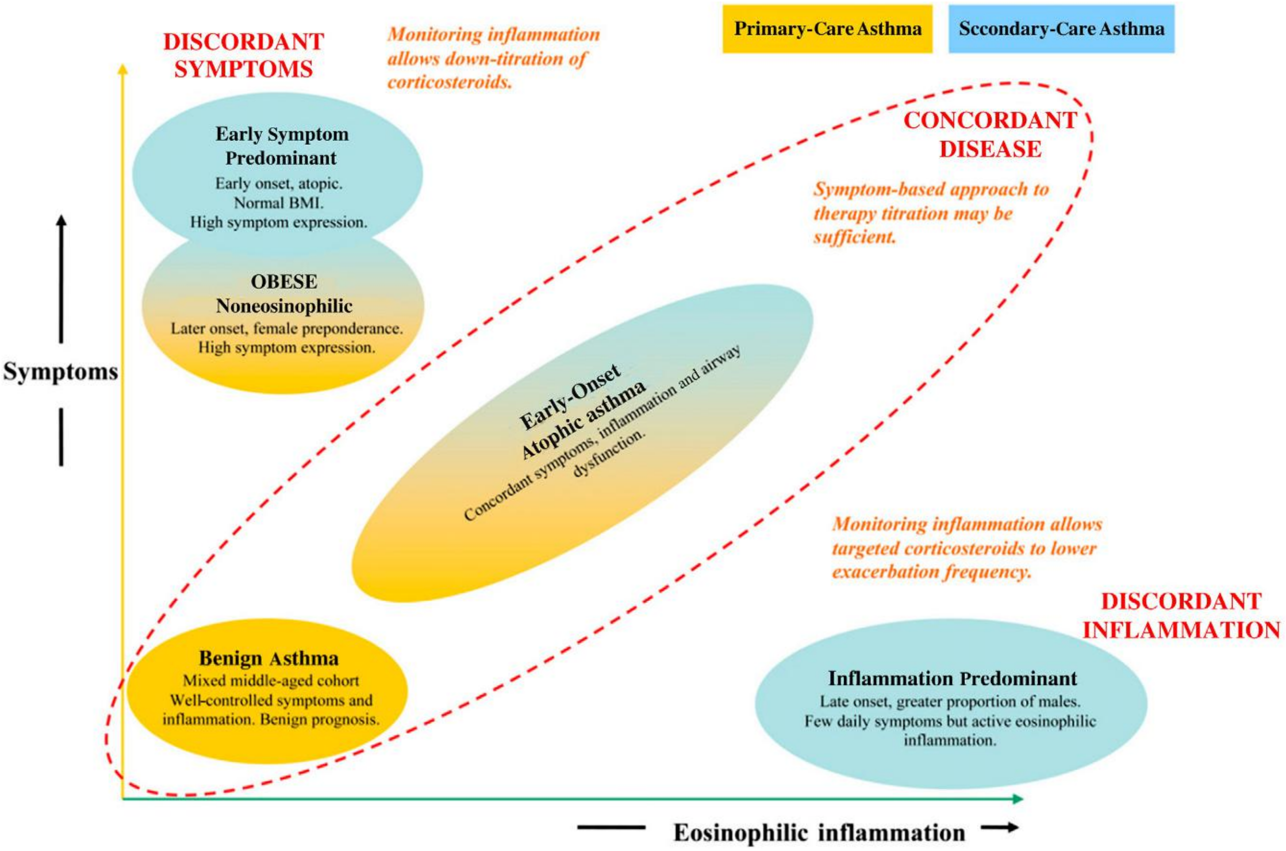

Primary-Care Asthma
Sccondary-Care Asthma
DISCORDANT SYMPTOMS
Monitoring inflammation allows down-titration of corticosteroids.
Early Symptom Predominant
Early onset, atopic. Normal BMI. High symptom expression.
OBESE Noneosinophilic
Later onset, female preponderance. High symptom expression.
Symptoms
CONCORDANT DISEASE
Symptom-based approach to therapy titration may be sufficient.
Early-Onset Atophic asthma
Concordant symptoms, inflammation and airway dysfunction.
Monitoring inflammation allows targeted corticosteroids to lower exacerbation frequency.
DISCORDANT INFLAMMATION
Benign Asthma
Mixed middle-aged cohort Well-controlled symptoms and inflammation. Benign prognosis.
Inflammation Predominant
Late onset, greater proportion of males. Few daily symptoms but active eosinophilic inflammation.
Eosinophilic inflammation

FIGURE 28.5 **Clinical phenotypes of asthma identified by cluster analyses.** Summary of phenotypes identified using cluster analysis in primary- and secondary-care asthma populations. *Source: Reproduced with permission from Halder et al.* [60].

and intrinsic asthma, and that treatment with ICS was effective in the majority of mild to moderate asthma cases, the distinctions between extrinsic and intrinsic asthma fell out of favor [64].

28.5.1.3 Heterogeneity of Responses to Existing Therapies

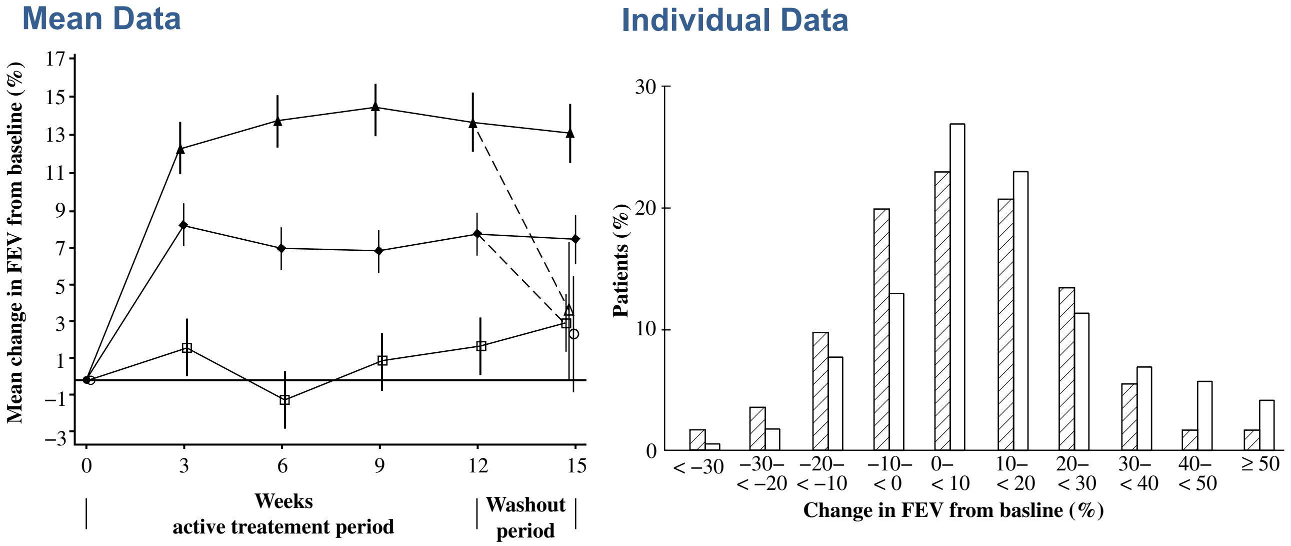
Heterogeneity in response to existing medication is exemplified by Malmstrom's study, which investigated individual response to leukotriene receptor antagonist (LTRA) (montelukast 10 mg once daily), and the inhaled steroid beclomethasone (200 µg twice daily) over a 12-week period in 895 chronic asthma patients [65]. Overall, there was a clear improvement in lung function (primary outcome FEV_1) for both treatments; however, when stratified based on individual patient data, there was a very large degree of heterogeneity for both treatment groups, with some patients showing up to 50% improvement in FEV_1 and others showing a decline in lung function with a 30% decrease in FEV_1 (Figure 28.6). This interindividual variability, a component of clinical trials that is seldom appreciated, is thought to have a strong genetic component.

28.5.2 Problems with Current Therapy

The therapeutic regimen for the treatment of asthma has been well defined and was described earlier in this chapter. Despite such well-defined treatment regimens, which are constantly updated and streamlined, as well as the availability of high-quality medications, in many cases asthma is still not sufficiently controlled. While current asthma therapeutics are able to successfully treat 90–95% of the asthmatic population, around 50% still have daily symptoms and almost all patients report limitations of daily activities [66]. This also leaves around 5–10% of patients who do not respond to conventional treatments. This demonstrates that the current one-size-fits-all approach is not perfect, and emphasizes the need for a stratified medicine approach to asthma.

28.5.2.1 Variable Efficacy and Presentation of Side Effects

Variability in current therapy involves variation in drug efficacy and presentation of side effects, each of which presents challenges to successful maintenance of asthma.


Mean Data

Individual Data

FIGURE 28.6 **Heterogeneity in asthma patient responses to current asthma medication.** While mean data demonstrates a significant improvement in lung function (FEV_1) over 12 weeks for both montelukast (10 mg once daily; circle) and beclomethasone (200 μg twice daily; triangle) versus placebo (open square) in 895 chronic asthma patients, individual data show a broad distribution of responses (Striped bars represent patients receiving montelukast, 10 mg once daily; white bars represent patients receiving inhaled beclomethasone, 200 mg twice daily). *Source: Reproduced with permission from Malmstrom et al.* [65].

Variability in drug efficacy, especially when occurring for β_2-adrenergic receptor agonists and glucocorticosteroids, and in the presentation of side effects, can result in patient therapy having to deviate from the standard therapeutic regimen for the treatment of asthma.

This high degree of interpatient variability in treatment response and in side effects can be attributed to a number of factors: (1) the interaction of genetic factors; (2) individual patient characteristics, such as weight, gender, and pregnancy; and (3) exposure to environmental insults, such as air pollution, allergens, and cigarette smoke. These factors interact, making the control of asthmatic episodes and the reduction of exacerbations more difficult. For example, asthma attacks can affect different age groups differently according to the season—children are more affected in summer while the older population suffers more in winter. When considering variation in treatment efficacy, one must also be mindful of individual patient factors such as inhaler technique and noncompliance (discussed earlier).

Currently asthma therapy is planned without consideration of genetic variance. However, genetic variance is arguably the most important of the factors listed previously—a claim supported by evidence gained from investigating the repeatability of response to therapy. This approach demonstrated that repeatability was between 60–80%, with a substantial proportion of the variance due to genetic factors [67].

28.5.2.2 Nonresponders

One important form of variance in therapeutic efficacy is non-response. Non-response is an important limitation in the maintenance and control of asthma, in that therapeutics expected to manage the disease at a certain level of severity fail outright, leading to the use of a trial approach in which different therapies are utilized with variable results until a suitable regimen is determined. A good example is severe asthma, where the patient is on permanent oral glucocorticosteroid therapy because of nonresponse to other therapies such as ICS, leukotriene receptor antagonists, and β_2-adrenergic receptor agonists. Non-responders present with longer duration of asthma exacerbations and worse morning lung function, as well as a more frequent family history of asthma, when compared to those individuals responsive to glucocorticosteroid therapy [67].

28.5.3 Nongenetic Approaches to Patient Stratification

28.5.3.1 Anti-IgE

The recognition that allergen specific IgE activation of mast cells is central in driving allergic asthma lead to the discovery that the primary mast cell-signaling cascade could be inhibited by a monoclonal antibody toward the IgE binding site to the high affinity receptor (FcER1). Anti-IgE (omalizumab) became the first specific biologic to be used in the treatment of severe allergic asthma. Clinical trials revealed efficacy as well as almost total inhibition of early and late asthmatic responses to inhaled allergen [68].

Following a rigorous examination of trial data by the National Institute for Clinical Excellence (NICE), omalizumab is now recommended as an option for treating severe persistent, confirmed allergic IgE-mediated asthma as an add-on to optimized standard therapy in people aged six and older who need continuous or frequent treatment with oral corticosteroids (defined as four or more courses in the previous year) in the United Kingdom. Although the drug is prescribed for cases of therapy-resistant asthma associated with allergy, it is also licensed for patients with evidence of allergy but a normal IgE. Omalizumab represents the first stratified treatment for asthma, but there is evidence that its use can be further stratified by using biomarkers of Th2 inflammation to predict response. In a recent paper studying the effect of omalizumab on uncontrolled severe persistent allergic asthma, treatment effect was analyzed in relation to F_ENO, blood eosinophils, and serum periostin (an epithelial protein that is induced by IL13). After 48 weeks of omalizumab, reductions in protocol-defined exacerbations were greater in high versus low subgroups for all three biomarkers, suggesting a potential prognostic ability [69].

28.5.3.2 Corticosteroids and Anti-IL13

This biomarker approach has helped divide asthma into and Th2 high and Th2 low disease. Asthma has traditionally been considered a Th2 process linked to atopy and allergy, type I hyper-sensitivity reactions, eosinophilic inflammation, and response to corticosteroids. "Th2 high disease," as evidenced by high F_ENO and a sputum eosinophilia of >2%, has been associated with a better response to corticosteroids when compared to "Th2 low disease" [70–72]. Studies on lebrikizumab, a novel monoclonal antibody to IL13 [73], have suggested that serum periostin may be a biomarker for a more general Th2 asthmatic phenotype. In one study a subgroup of individuals who had asthma and persistent elevation in serum periostin showed greater improvements in airway function and fewer exacerbations after lebrikizumab treatment than those with lower serum periostin. Interestingly, levels of F_ENO, produced by inducible nitric oxide synthase, an enzyme induced in human airway epithelial cells by IL13, were as informative as periostin in identifying Th2-high individuals responsive to lebrikizumab.

28.5.3.3 Anti-IL5

Both the role of stratification based on phenotyping and the need for correct study end points in treatment trials have been informed by the salutary lesson of the anti-IL5 antibody mepolizumab. Although mepoluzimab effectively blocks eosinophilic inflammation, initial studies were disappointing in that no effect was demonstrated on spirometry or peak flow, despite a significant reduction in blood eosinophil numbers [74]. When the effect of the drug was studied on an outcome measure related to eosinophilia, namely severe exacerbations, there were significantly fewer exacerbations with a concomitant improvement in asthma-related quality of life [75]. Tellingly, there were no significant differences with respect to symptoms, postbronchodilator FEV_1, or airway hyper-responsiveness, suggesting that different physiological processes determine these asthma outcomes. The study also helps researchers reflect upon the need to measure the correct end point in a disease that has different pathological mechanisms that all respond differently to treatment.

Phenotyping asthma has led to the identification of new disease targets, but aside from the use of anti-IgE, it has not yet led to a step change in management. Other approaches assessing the separate aspects of asthma pathophysiology—namely, airway inflammation, airway hyper-responsiveness, and airflow obstruction have been trialled with differing results. The key determinant in the design of these stratified approaches was the recognition that although asthma is a disease defined by symptoms, with treatment response also assessed by symptom improvement, there is no clear correlation between pathophysiology and symptomatology/exacerbation risk.

In the FACET study (designed to evaluate the benefits of adding a long-acting β_2-adrenergic receptor agonist to different doses of ICS), higher-dose ICS had a marked beneficial effect on exacerbation frequency, but relatively less effect on symptoms and peak expiratory flow. The opposite was true with the addition of long-acting β_2-agonists [76]. This indicates that exacerbation frequency does not closely relate to symptoms and measures of disordered airway function, suggesting that the mechanisms responsible for these features are different [77].

Numerous studies have demonstrated that in asthma the features of airway inflammation, airway hyper-responsiveness, variable airflow obstruction and associated symptoms can overlap, occur independently or change over time, in response to treatment or other external factors such as allergen exposure, or viral infection. One obvious example of this is eosinophilic bronchitis, a condition characterized by corticosteroid-responsive cough and the presence of a sputum eosinophilia occurring in the absence of variable airflow obstruction or airway hyper-responsiveness [78].

Although asthma guidelines recommend the assessment of airflow obstruction (FEV_1/PEF) and related symptoms for their primary treatment response outcomes, the realization is that this approach looks only at one aspect of asthma pathophysiology which is not closely related to exacerbation risk. This has led to new approaches aimed at reducing asthma exacerbations and symptoms while maintaining corticosteroid burden at the lowest possible level.

28.5.4 Clinicopathophysiological Measures in Asthma with Potential Use in Stratification

28.5.4.1 Airway Hyper-Responsiveness

Airway hyper-responsiveness is defined as increased sensitivity to an inhaled constrictor agonist and a steeper slope of the dose–response curve. Two main forms of bronchoconstrictor stimuli exist: direct and indirect. Direct bronchoconstrictors, such as histamine or methacholine, stimulate receptors on the airway smooth-muscle, while indirect ones cause bronchoconstriction by secondary release of bronchoconstrictor mediators from mast cells or activation of neural pathways. Airway responsiveness is usually measured as the provocative dose of methacholine causing a 20% fall in FEV_1 by linear interpolation of the log dose–response curve (PC_{20}). In the general population, the distribution of airway hyper-responsiveness follows a continuous unimodal log-normal distribution, with asthma sufferers representing the hyper-responsive part of the distribution curve. A PC_{20} is not usually measurable in non-diseased individuals, which suggests a large difference in airway responsiveness between non-diseased individuals and asthma patients.

The cut-off used to identify asthma is normally a methacholine concentration of <8 mg/ml. This value had a sensitivity of 100%, a specificity of 93%, and a negative predictive value of 100% in a study on a population of 500 college students with a diagnosis of current symptomatic asthma. The use of methacholine PC_{20} to diagnose asthma has been evaluated; one study demonstrated that when asthma is defined as consistent symptoms with objective evidence of abnormal variable airflow obstruction, a positive methacholine challenge is more sensitive than PEF amplitude % mean and the acute bronchodilator response in diagnosis [79].

The use of methacholine PC_{20} to guide treatment has also been assessed. Although it was found that PC_{20}-guided treatment resulted in a reduction in asthma exacerbations, this was at the expense of increased inhaled corticosteroid use [80]; consequently, although PC_{20} is often used for asthma diagnosis, or as a study end point, it is not routinely used to guide treatment decisions.

28.5.4.2 *Airway Inflammation*

Asthma has been traditionally viewed as a condition where airway inflammation causes airway hyper-responsiveness, which in turn leads to variable airflow obstruction and symptoms; however, cross-sectional and longitudinal studies of airway inflammation using sputum induction in large populations with a diverse range of presentations suggest that this hypothesis requires modification. A recent bronchoscopy study demonstrated that bronchoconstriction is independent of airway inflammation and can lead to airway remodeling [38].

The development of noninvasive techniques to assess airway inflammation, including induced sputum and F_ENO, has made it possible to relate airway inflammation to objective measures of disordered airway function in larger and more heterogeneous populations than was possible with bronchoscopy studies. In general, these newer studies contradict findings in earlier studies and do not find a correlation between sputum eosinophil count and various markers of airway dysfunction.

It has been observed that a subset of patients with symptomatic asthma do not have sputum evidence of eosinophilic airway inflammation [81]. Many have sputum neutrophilia. This sputum profile is evident in corticosteroid-naïve as well as corticosteroid-treated subjects, suggesting that it is not always an artifact related to treatment. Importantly, patients with noneosinophilic asthma respond less well to inhaled budesonide than do a group with more typical sputum features [82].

Similar sputum findings have been reported in more severe asthmatics; a subgroup of patients with refractory asthma have been identified who have bronchoscopic evidence of neutrophilic airway inflammation, normal eosinophil counts, and a normal basement membrane thickness. These findings suggest the presence of a distinct asthma phenotype characterized by a predominantly neutrophilic airway inflammatory response and relative corticosteroid resistance. Furthermore, there is evidence that neutrophilic asthma may result from activation of the innate immune system with the production of proinflammatory cytokines.

The use of noninvasive measures to guide treatment decisions by stratification based on the presence or absence of airway inflammation has been assessed. The strongest evidence for effect relates to the use of induced sputum to guide corticosteroid dose in moderate to severe asthma. Two studies have shown that using induced sputum differential counts to guide treatment results in fewer exacerbations for the same overall corticosteroid burden [81,83].

Although the evidence is clear that this approach works, induced sputum is not widely employed, possibly because it is seen as time consuming, requiring expertise both to induce the sputum and to perform the differential counts. The discovery that levels of F_ENO, measured via a device similar to a breathalyzer, correlate well with sputum eosinophilia and relate to corticosteroid responsiveness, has driven the application of inflammometry (noninvasive measurement of inflammation in the airways). Inflammometry using F_ENO measurements has also been used to guide and stratify treatment decisions. Though one study was positive in patients with asthma in pregnancy, with a reduction in exacerbations, most have failed to show an improvement when compared to guideline-driven (FEV_1/PEF and symptoms) management. The hunt is still on for a reliable, inexpensive, and valid biomarker of airway inflammation.

28.5.4.3 Small Airways

FEV_1 and PEF measurements reflect changes in the caliber of the large airways. Our knowledge of anatomical and physiological changes in the small airways of patients with asthma is based on small case series of resected lung tissue from patients with asthma undergoing surgery for cancer, or on cases of fatal asthma.

These case series have demonstrated that there is significant inflammation present in the small airways (<2 mm diameter) in asthma. Fatal asthma is associated with peripheral airway inflammation and differences in the number of activated eosinophils in the distal lung. Other studies have revealed alterations in the epithelium and smooth-muscle, as well as mucous hypersecretion and distal airway plugging of the small airways. The presence of inflammation in the small airways in asthma may explain why small airways account for up to 50–90% of total airflow resistance in asthma, but only 10% of airflow resistance in normal airways.

Recently, the development of "small-particle" ICS, designed to target the peripheral lung, and the advent of new technologies—nitrogen washout, impulse oscillometry, and hyperpolarized noble gas magnetic resonance imaging, which allows assessment of peripheral lung function—have led to a resurgence of interest in the distal lung. Studies of small-particle ICS have been inconsistent; those comparing small-particle and standard-particle ICSs have failed to demonstrate improved asthma outcomes when administered in clinically comparable doses. Future asthma treatment may yet be stratified by the presence or absence of small airway inflammation.

28.5.4.4 Stratifying Treatment Response in Children Based on Clinical Parameters

There have been several attempts to base treatment decisions on measures of airway inflammation, including induced sputum and F_ENO in children. Generally, they have been unsuccessful, although some experts do stratify treatment decisions on the response to oral or intramuscular steroids.

28.5.5 Current Asthma Therapies—Pharmacogenetic Approaches to Stratification

While methods of stratifying asthma patients to specific treatments based on nongenetic factors such as clinical outcomes, cellular measures, or protein biomarkers have shown some success, a large body of work has investigated the potential of genetic markers as predictors of patient responses to existing therapies, i.e., pharmacogenetics. Pharmacogenetics, the investigation of the effect of genetic polymorphisms on response to treatment or risk of adverse side effects, is one of the first steps in developing personalized prescribing. The use of genetic information to stratify patient prescribing is potentially more desirable compared to nongenetic stratification, as technology now allows a small amount of blood or saliva sample to be taken and a DNA test can be completed within hours.

To date, asthma pharmacogenetic studies have suffered from relatively small retrospective designs and a focus on only a few candidate genes; however, more recent, larger prospective studies have been completed that provide greater confidence in original findings and hypothesis-free approaches such as GWAS. These studies are primarily focused on pharmacodynamic aspects (e.g., improvement in lung function post-treatment), and only limited information is available about the impact of pharmacogenetics on adverse effects.

Pharmacogenetics in asthma is relatively advanced compared to that for other diseases. Genetic factors influencing the main treatment classes (i.e., β_2-adrenergic receptor agonists, corticosteroids, and leukotriene modifiers) have been identified with some confidence (Table 28.3). An overview of these findings follows; however, for more in-depth analyses, see Portelli and Sayers [84].

28.5.5.1 β_2-Adrenergic Receptor–Agonist Pharmacogenetics

β_2-adrenergic receptor agonists carry out their function through the β_2-adrenergic receptor—a 413-amino-acid G-protein-coupled receptor encoded by an intronless gene (*ADRB2*) located on chromosome 5q31.32. Binding of agonists to the receptor activates adenyl cyclase through stimulatory Gs proteins, which in turn activate protein kinase A. The latter phosphorylates several target proteins, resulting in a decrease in intracellular calcium that, importantly, causes smooth-muscle relaxation in the airways. It is not surprising that the majority of evidence of genetically driven effects on β_2-adrenergic receptor–agonist therapy stem from the *ADRB2* gene, which has therefore been extensively studied for pharmacogenetic effects on β_2-adrenergic receptor–agonist responses.

ADRB2 is a highly polymorphic gene containing 51 known and validated polymorphisms, of which 49 are SNPs and 2 are insertion/deletion variants. Most studies have focused on the role of four nonsynonymous coding-region polymorphisms: Arg16Gly (Arginine-to-Glycine substitution at position 16 in the protein), Gln27Glu, Val34Met, and Thr164Ile [85]. In a Caucasian population, the frequency of the polymorphisms at positions 16 and 27 were identified as 59% (Arg16Glu) and 29%, (Gln27Glu) [85]. The remaining Val34Met and Thr164Ile polymorphisms are rare, having approximate frequencies of <0.001% and 0.05%, respectively. The Arg16 variant has been shown to have pharmacogenetic potential through association with;

TABLE 28.3 Recently Identified Pharmacogenetic Effects on Current Asthma Therapies

Gene	Key Polymorphism	Pharmacogenetic Effect
β_2-adrenergic receptor agonists (e.g., salbutamol, salmeterol)		
ADRB2	Arg16Gly	Mixed reports (see text)
	Gln27Glu	Mixed reports (see text)
	PolyC (3′UTR)	Counteracts *ADRB2* downregulation
ARGI	rs2781659 (5′ region)	Haplotypes increase *ARG1* transcription, leading to improved response to β_2AR agonist
GSNOR	rs1154400 (5′ region)	Differential response to β_2AR agonist; *GSNOR* haplotypes determine *GSNOR* transcription
SPATS2L	rs295137 (5′ region)	Differential response to β_2AR agonist; *SPATS2L* expression inversely related to β_2AR levels
COL22A1	rs6988229 (intron)	Differential response to β_2AR agonist
Leukotriene modifiers (e.g., montelukast, zafirlukast, zileuton)		
ALOX5	rs2115189 (intronic)	Causes differential response to zileuton and montelukast therapy (FEV$_1$ change)
	rs4987105 (Thr120Thr)	Causes differential response to montelukast therapy (FEV$_1$ change)
	rs4986832 (5′ region)	
LTC4S	A-444C	Associated with LTRA and LTSI response (FEV$_1$ change) in some but not all studies.
	rs272431(intron)	Causes differential response to zileuton therapy (FEV$_1$ change)
CYSLT2	rs91227 (3′UTR) rs912278 (3′UTR)	Associated with morning PEF in subjects taking montelukast
MRP1	rs119774 (intronic)	Causes differential response to zileuton and montelukast therapy (FEV$_1$ change)
OATP2B1	rs12422149 (Arg312Gln)	In some but not all studies, reduced morning plasma concentrations of montelukast, resulting in lack of clinical benefit
CHRM2	4 SNPs	Causes differential response to montelukast therapy (FEV$_1$ change)
Corticosteroids (e.g., prednisolone, beclomethasone, fluticasone)		
CRHR1	rs242941 (intronic)	Associated with corticosteroid response (FEV$_1$ change)
TBX21	rs224001 (His33Gln)	Improvements in BHR with corticosteroid treatment
NK2R	rs7703891 (Gly231Glu)	Improved asthma control with corticosteroid treatment
STIP1	rs4980524 (intron 1) rs6591838 (intron 1) rs2236647 (intron 5)	Associated with corticosteroid response (FEV$_1$ change)

(*Continued*)

TABLE 28.3 Continued

Gene	Key Polymorphism	Pharmacogenetic Effect
DUSP1	rs881152 (5′ region)	Associated with bronchodilator response and asthma control in presence of corticosteroids
FCER2	rs28364072 (intron)	Associated with asthma exacerbations in presence of corticosteroids
GLCCI	rs37972 (5′ region)	Associated with corticosteroid response (FEV_1 change); potential transcriptional mechanism
CRHR1	5 SNPs	ACQ post-corticosteroid
COL2A1	3 SNPs	FEV_1 change post-corticosteroid
CYP3A4	CYP3A4*22	Asthma control post-corticosteroid; *CYP3A4*22* previously associated with *CYP3A4* levels
Combination Therapy (β_2-adrenergic receptor–agonist and corticosteroid)		
CHRM2	3 SNPs	ACQ score
HSPA8	rs1461496	FEV_1 change
NOS3	rs1799983 (Glu298Asp)	FEV_1 change

Note: See main text for abbreviations and key references. ACQ—Asthma Control Questionnaire; FEV_1—forced expiratory volume in 1 second.

- An enhanced acute response to β_2-adrenergic receptor agonists
- A decline of asthma control following prolonged use of β_2-adrenergic receptor agonists
- A subsensitivity of response for bronchoprotection by β_2-adrenergic receptor agonists

However, several studies have failed to reproduce these effects, meaning that a common consensus on the contribution of Arg16 has yet to be reached. This lack of consensus can be explained when we consider that there are at least 12 different haplotypes (a specific combination of SNPs across the gene) in *ADRB2*. Investigations into the effect of a genotype in isolation, without consideration of its haplotype, is likely to introduce confounding into the association.

Functional effects of coding-region polymorphisms in *ADRB2* have been identified through *in vitro* work carried out in cell lines. These have included the following:

- Impaired agonist binding and reduced adenyl cyclase activity in the Thr164Ile variant receptor
- An enhanced agonist-mediated receptor downregulation for the Gly16 variant
- A resistance to downregulation for the Glu27 variant

More recent clinical studies have been hindered by relatively small sample sizes. For example, Lee et al. identified an association between Gln27Glu and Terbutaline nebulizer response and an absence of effect for the Arg16Gly polymorphism in children, in 27 patients [86]. In one of the larger studies, Basu et al. identified an Arg16 copy number-dependent increase in disease exacerbations in 1182 patients with asthma aged 3 to 22 years on daily exposure to β_2-adrenergic receptor agonist (regularly inhaled corticosteroid plus salbutamol on demand group, and regularly inhaled corticosteroid plus salmeterol and salbutamol on demand group) [87].

It is important to note that this effect was driven by asthma patients who used salbutamol and/or salmeterol daily, supporting the suggestion that the Arg16 polymorphism has an integral role in the effectiveness of β_2-adrenergic receptor–agonist therapy [87]. However, as with other studies, study limitations did not allow a clear association to be made between the *ADRB2* gene (via its polymorphisms) and β_2-adrenergic receptor–agonist efficacy. Namely, all participants were using β_2-adrenergic receptor–agonists as a reliever; a study arm of non-β_2-adrenergic receptor–agonist reliever use would have provided a clearer interpretation of the detrimental effects of salbutamol/salmeterol in the Arg16 subjects.

Large studies have failed to observe a clinically relevant effect of these polymorphisms. These include longitudinal studies of β_2-adrenergic receptor–agonist efficacy and the study of concomitant administration of corticosteroids.

A good example is a study of 2250 asthma patients randomly assigned to: (1) budesonide plus formoterol maintenance and reliever therapy; (2) fixed-dose budesonide plus formoterol; or (3) fixed dose fluticasone plus salmeterol for six months. No overall effect of the Gly16Arg genotype on clinical outcomes was found [88]. Another contributing factor to the pronounced differences in the conclusions from different studies investigating these *ADRB2* polymorphisms is trial design variation.

Other *ADRB2* Polymorphisms

The majority of studies have focused on the *ADRB2* Arg16Gly polymorphism; however, other potential pharmacogenetically relevant polymorphisms also affect β_2-adrenergic receptor–agonist efficacy. Multiple polymorphisms in the gene's regulatory regions that have potential clinical relevance are present in *ADRB2*. By studying eight common haplotypes based on 26 SNPs, a recent *in vitro* approach using "whole gene" transfection identified differential effects on receptor expression and downregulation that are haplotype-driven [89]. This study identified four common haplotypes with elevated receptor expression and two haplotypes with enhanced receptor downregulation. Another area recently investigated is polymorphisms occurring in the gene's untranslated regions, which may have an effect on gene expression and resultant drug efficacy.

28.5.5.2 Novel Regulators of β_2-Adrenergic Receptor–Agonist Responses

Multiple genes are likely to be involved in the regulation of β_2-adrenergic receptor–agonist response and expression of side effects, due to the agonists' known complex and multifactorial mechanism of action. One gene that has been associated with patient response to these agonists is Arginase 1 (*ARG1*), which was identified using a novel algorithm implemented in a family-based association test (FBAT) [90]. In this study of 209 children and their parents, the *ARG1* SNP, rs2781659, was associated with bronchodilator response (BDR) when SNPs from 111 candidate genes (42 involved in β_2-adrenergic receptor–signaling/regulation, 28 involved in glucocorticoid regulation, and 41 from prior asthma association studies) were investigated for their association with acute response to inhaled β_2-adrenergic receptor–agonist in 209 children and their parents. In agreement with this study polymorphisms spanning *ARG1* and influencing patient response to salbutamol have also been identified in a candidate gene study involving 221 asthma subjects [91].

The *ARG1* polymorphisms identified in both studies were in linkage disequilibrium (LD) (inherited together), suggesting a common causative mechanism involving potential transcriptional regulation due to the polymorphisms' location (predominantly 5′ to the gene). This alteration in transcription has now been confirmed in promoter-reporter studies, which found that the key *ARG1* haplotype associated with improved BDR drives the highest level of *ARG1* promoter activity [92]. In the study by Vonk et al., *ARG2* SNPs were also associated with patient responses to salbutamol [91], suggesting an integral role for the arginase family in β_2-adrenergic receptor–agonist therapy.

Recently a S-nitrosoglutathione reductase (*GSNOR*) SNP (rs1154400, promoter region) was associated with a decreased response to salbutamol in 107 African-American children [93]. In the same study, a post hoc multilocus analysis discovered that a combination of rs1154400 with *ADRB2* Arg16Gly, Gly27Glu, and the carbamoyl phosphate synthetase-1 (*CPS1*) SNP rs2230739 gave a 70% predictive value for lack of response to therapy [93], implying that pharmacogenetic regulation of β_2-adrenergic receptor–agonist therapy may depend on several loci acting together via gene–gene interactions. In confirmation, 4/5 SNPs tested in *GSNOR* were associated with asthma patient responses to salbutamol in 168 Puerto Rican asthma patients [94]. These SNPs were also associated with asthma susceptibility, and the key risk haplotype was associated with increased transcriptional activity based on promoter-reporter studies [94]. *GSNOR* is an alcohol dehydrogenase that breaks down GSNO, an endogenous bronchodilator [95]. In addition, GSNO regulates nitrosylation of proteins, leading to alterations in function, including G-protein coupled receptor kinase 2 (*GRK2*), which phosphorylates and desensitizes the β_2-adrenergic receptor [96].

Other novel genes have recently been associated with β_2-adrenergic receptor–agonist therapy including the spermatogenesis–associated, serine-rich 2-Like (*SPATS2L*) and collagen (*COL22A1*) genes [97]. The *COL22A1* gene, through association with the intronic SNP rs6988229, was associated with acute bronchodilator response to inhaled salbutamol in a genome-wide association study in which ~500,000 SNPs were tested in 403 Caucasian trios. This association was replicated in a pooled population of three asthma trial populations, as well as in three additional asthma populations [97]. *SPATS2L* was also identified as a modulator of β_2-adrenergic receptor–agonist function through a GWAS. Here, the SNP rs295137, located near the *SPATS2L* gene, was significantly associated with percentage change in baseline FEV_1 in 1644 Caucasian asthma patients; this was replicated in two alternate Caucasian populations ($n \sim 500$ each). Molecular biology techniques confirmed these results and identified that *SPATS2L* may be an important regulator of β_2-adrenergic receptor downregulation.

The identification of patients at risk from potential adverse effects of β_2-adrenergic receptor–agonists remains a critical clinical question, as does the targeting of this class of drug to patients most likely to benefit from it. While there has been clear progress with respect to study design (e.g., examining haplotypes instead of genotypes in isolation) and adequately powered studies using thousands of individuals, there is still a need for large prospective studies of asthma patients with matched phenotypes and carefully controlled covariates that include environmental influences. Similarly, these studies would benefit from GWA approaches and the identification of gene–gene interactions.

28.5.6 Pharmacogenetic Approaches to Disease Management: *ADRB2* Arg16 and Second-Line Therapy

While investigations of the effect of genetic polymorphisms on β_2-adrenergic receptor–agonist responses have predominantly focused on clinical end points (e.g., lung function parameters) in the different genotype groups, there has been recent interest in the real-life application of genetic knowledge. As outlined, the first-line treatment for asthma is a short-acting β_2-adrenergic receptor–agonist (e.g., salbutamol) as needed (step 1); if symptoms persist, the addition of inhaled corticosteroid (e.g., beclomethasone) is considered (step 2); for further control, a long-acting β_2-adrenergic receptor agonist (e.g., salmeterol) or a leukotriene receptor antagonist (LTRA) (e.g., montelukast) is added (step 3) (Figure 28.4).

In a recent study, Lipworth et al. set out to use genetic information on the *ADRB2* Arg16 polymorphism to inform the choice of prescribing salmeterol or montelukast as add-on therapy [98]. Children with persistent asthma and homozygous for the Arg16 genotype ($n = 62$) were randomized to receive salmeterol (50 μg, bd) or montelukast (5 or 10 mg, once daily) as an add-on to inhaled fluticasone propionate for one year. The study tested whether carriers of the Arg16 genotype were more prone to adverse effects (e.g., prolonged β_2-adrenergic receptor–agonist use associated exacerbations), and hence whether montelukast provided superior control for this preselected population.

Outcomes were school absences (primary outcome), exacerbation score, reliever use (salbutamol), morning dyspnoea, and Asthma Control Questionnaire (ACQ) quality-of-life scores. Montelukast provided superior benefit for all measures, with clinically relevant differences within three months. No significant difference in FEV_1 (% Pred) was observed, providing further evidence of limited correlation between lung function and symptom-based scores [98]. These results suggest that larger and longer prospective studies are warranted to provide more definitive data on the clinical utility of Arg16 stratification including a Gly16 study arm and the use of additional markers to define the population would provide clearer interpretation.

28.5.7 Leukotriene Synthesis Inhibitor and Leukotriene Receptor Antagonist Pharmacogenetics

As outlined, leukotriene synthesis inhibitors (LTSIs) and leukotriene receptor antagonists (LTRAs) are commonly used as an add-on therapy in asthma to provide greater control or steroid-sparing effects. The cysteinyl leukotrienes (LTC_4, LTD_4, LTE_4) and dihydroxy leukotriene (LTB_4) contribute to the inflammatory process in asthma and are synthesized from arachidonic acid via the 5-lipoxygenase pathway (Figure 28.7). Cysteinyl leukotrienes have been implicated in bronchoconstriction, mucus secretion, vascular permeability, inflammatory cell infiltration and cytokine production. Arachidonic acid is converted to 5-hydroperoxyeicosatetraenoic acid and leukotriene A_4 (LTA_4) by 5-lipoxygenase (5-LO/*ALOX5*) and 5-lipoxygenase-activating protein (FLAP/*ALOX5AP*), which acts as an adaptor protein for this reaction. LTA_4 is then converted to leukotriene B_4 (LTB_4) by LTA_4 hydrolase (LTA4H) or, alternatively, is conjugated with reduced glutathione to form LTC_4 via the actions of leukotriene C4 (LTC_4) synthase (LTC4S). LTC_4 is next transported to the extracellular space via the multidrug resistance protein 1 (MRP1) (Figure 28.7). Leukotriene modifiers include LTSIs, which act by targeting 5-LO (e.g., zileuton), resulting in a decrease in all leukotriene biosynthesis (LTC_4, LTD_4, LTE_4, and LTB_4) or LTRAs, which act by specifically blocking cysteinyl leukotrienes from binding to their primary receptor, CYSLTR1, which is found on many cell types including inflammatory cells and airway smooth-muscle cells, (e.g., montelukast and zafirlukast).

Data demonstrating the large degree of heterogeneity of asthma patient responses to this class of drug are summarized in Figure 28.6. These data have led to extensive genetic studies investigating the roles of SNPs in a large number of candidate genes associated with leukotriene production and/or activity and acute responses to LTRAs and LTSIs. Associations have been described for polymorphisms in *ALOX5*, *ALOX5AP*, *LTC4S*, *CYSLTR1*, *CYSLTR2*, and *MRP1*; however, many of these studies have been small and so have led to a lack of reproducibility in findings [99]. More recent studies have looked at different aspects of leukotriene modifier functions, including the association of montelukast absorption (i.e., pharmacokinetics) with polymorphisms in *OATP2PB1* [100]. However, others have failed to observe these associations [101].

28.5.7.1 5-Lipoxygenase (ALOX5)

5-LO (along with 5-LO-activating protein) is the major regulatory switch for leukotriene production, and extensive studies have demonstrated that the level of 5-LO can have dramatic effects on both cysteinyl and dihydroxy leukotriene production. 5-LO catalyzes the conversion of arachidonic acid to LTA_4, one of the early stages in leukotriene production (Figure 28.7). *ALOX5* is found on chromosome 10q11.2 and is a large gene composed of 14 exons and 13 introns; it spans approximately 82 kb. Multiple studies suggest that *ALOX5* polymorphisms can influence clinical responses to LTRAs and LTSIs—in particular, a functional repeat polymorphism in the promoter region (resulting in alterations in SP1 transcription factor binding) have been associated with altered gene transcription and with response to the LTSI, ABT-761 [99,102].

These early studies have now been extended to investigate multiple polymorphisms, spanning the entire gene, and association with LTRA and/or LTSI responses in asthma patients. For example, Tantisira et al. observed an association between *ALOX5* intronic SNPs—rs892690, rs2029253, and rs2115819—and change in FEV$_1$ post-LTSI(zileuton)—in a cohort of 577 asthma patients [103]. The rs2115819 SNP was also a predictor of response to the LTRA, montelukast, in a previous study of 252 asthma subjects using percentage change in FEV$_1$ as the primary outcome [104]. In both cases, the GG versus GA or AA genotype had the greatest improvement in lung function (FEV$_1$) post-drug [104]. The lack of association for the rs892690, and rs2029253 SNPs in the montelukast study may be due to the reduced power of this study compared to the zileuton study. The functional significance of these SNPs remains to be resolved, and it is important to note that they may not be the causative genetic change, and that polymorphisms inherited at the same time (i.e., in LD) may be of relevance.

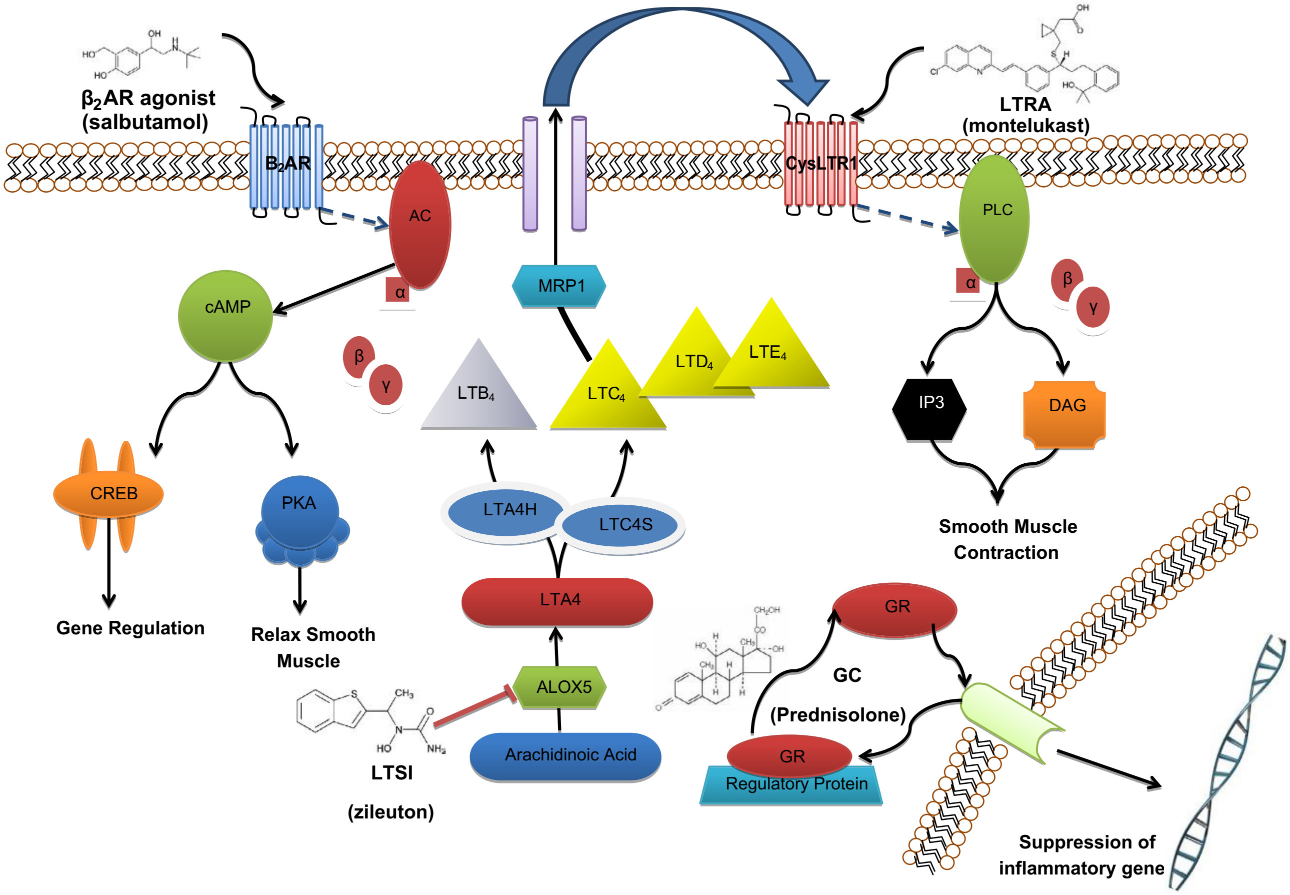


FIGURE 28.7 Targets of current asthma medication. β$_2$-adrenergic receptor agonists bind to the β$_2$-adrenergic receptor and mediate smooth-muscle relaxation via cAMP production and PKA, and regulate various genes via the transcription factor CREB. The cysteinyl leukotrienes LTC$_4$, LTD$_4$ and LTE$_{4,}$ and dihydroxy leukotriene, LTB$_4$ are produced in a series of steps via an enzymatic cascade. Following translocation into the extracellular space, cysteinyl leukotrienes activate their cognate GPCRs (including CysLTR1) and mediate a variety of effects, including smooth-muscle contraction. Corticosteroids enter the cell and bind to receptors that are stabilized by regulatory proteins. The glucocorticoid receptor (GR) then crosses the nuclear membrane, binds to a response element, and modulates transcription. AC—adenyl cyclase; ALOX5—arachidonate 5-lipoxygenase; β2AR agonist—beta2-adrenergic receptor agonist; β2AR—beta2-adrenergic receptor; cAMP—cyclic adenosine monophosphate; cysLTR—cysteinyl leukotriene receptor; CREB—cyclic AMP response-element binding; DAG—Di-acyl Glycerol; GC—glucocorticoid; GPCR—G-protein-coupled receptor; GR—glucocorticoid receptor; IP3—inositol-tri-phosphate; LTA$_4$—leukotriene A$_4$; LTB$_4$—leukotriene B$_4$; LTC$_4$—leukotriene C$_4$; LTD$_4$—leukotriene D$_4$; LTE$_4$—leukotriene E$_4$; LTC4S—leukotriene C$_4$ synthase; LTA4H—leukotriene A$_4$ hydrolase—MRP1—multidrug resistance protein 1; LTSI—leukotriene synthesis inhibitor; LTRA—leukotriene receptor antagonist; PKA—cyclic AMP-dependent protein kinase; PKC—protein kinase C; PLC—phospholipase C. *Source: Reproduced with permission from Portelli and Sayers* [84].

Two *ALOX5* SNPs (rs4987105 (synonymous Thr120Thr) and rs4986832 (5′ region)) were also associated with improvement in PEF on treatment with montelukast in a 12-week study in 174 asthma patients. This further confirmed the relevance of *ALOX5* SNPs and this key enzyme in leukotriene production [105]. Additional support for the relevance of *ALOX5* in asthma control is provided by a recent study of 270 children which found that the *ALOX5* Sp1 promoter polymorphism determined urinary LTE_4 levels and was associated with reduced lung function and, potentially, an ACQ score in non-5 copy carriers [106]. In this study, there was an inverse relationship between urinary LTE_4 levels and FEV_1.

28.5.7.2 Leukotriene C_4 Synthase

Another important enzyme in the leukotriene pathway is leukotriene C_4 synthase (*LTC4S*), which specifically results in the generation of the first cysteinyl leukotriene, LTC_4, by conjugating glutathione to LTA_4. There has been intense research into the role of *LTC4S* in asthma susceptibility because this enzyme, like 5-LO, is thought to be a key regulatory switch for cysteinyl leukotrienes, which are thought to play a more prominent role in asthma (e.g., via bronchoconstriction) than the dihydroxy LTB_4. *LTC4S* is relatively small, spanning 2.51 kb on chromosome 5q35. In particular, a promoter polymorphism (A-444C, rs730012) has been intensely investigated for pharmacogenetic effects as it is thought to alter *LTC4S* levels via transcription (higher levels in C allele carriers) and therefore alter LTC_4 production [107].

Of the ten published studies investigating the role of *LTC4S*-444 A>C, five showed an improvement in outcome measures, including FEV_1 post-LTRA therapy for carriers of the C allele, as hypothesized, although results were not always statistically significant. In a more recent study of zileuton responses (FEV_1 change over time) in 577 asthma subjects, the *LTC4S*-444 polymorphism showed no effect; however, an alternative *LTC4S* polymorphism (rs272431, intron 1) was associated with improved response (mean FEV_1) [103]. As these results were obtained in different populations and using different end points, the relative contribution of the *LTC4S* polymorphism to LTRA or LTSI therapeutic responses in asthma are still unclear. Larger prospective studies using multiple ethnic groups are required.

28.5.7.3 Cysteinyl Leukotriene Receptor 1 and 2 Genes (CYSLTR1/CYSLTR2)

Montelukast and other LTRAs target the cysteinyl leukotriene receptor 1, which is expressed in a variety of cells, including airway smooth-muscle and various inflammatory cells (e.g., eosinophils), therefore inhibiting the activation of this receptor by cysteinyl leukotrienes (Figure 28.7). Cysteinyl leukotriene receptor 1 is thought to be the main receptor mediating cysteinyl leukotriene receptor smooth-muscle contraction and inflammatory cell cytokine production in asthma. Again, genetic variation in the target receptor for these compounds may influence how effective they are in carriers of these alleles by regulating receptor function/expression. The cysteinyl leukotriene receptor 1 gene, *CYSLTR1*, is intronless and found on chromosome Xq13-21; it generates a 337 amino acid G-protein-coupled receptor (GPCR).

A second receptor for cysteinyl leukotrienes has also been described and, while not the target of LTRAs, may modulate the effects of cysteinyl leukotrienes *in vivo*. The gene for cysteinyl leukotriene receptor 2 is found on chromosome 13q14 and encodes for a 346 amino acid GPCR. Interestingly, in combination with several *ALOX5* SNPs, polymorphisms spanning *CYSLTR1* have been identified as determinants of responses to montelukast, suggesting SNP–SNP interactions may be important [108]. However, direct evaluation of *CYSLTR1* polymorphisms and responses to montelukast [104] or zileuton [102] have not identified a significant effect for them to date. Studies of *CYSLTR2* are limited; however, SNPs rs91227 and rs912278 (3′UTR) have been associated with an improvement in morning PEF following administration of montelukast for 12 weeks [105].

28.5.7.4 Novel Genes Associated with Clinical Responses to Leukotriene Modifier Drugs

While the majority of studies have focused on candidate polymorphism analyses in genes directly involved in leukotriene production or activity, several studies have identified additional genes of relevance to LTSI and LTRA response. In addition to its effects on vasodilation and bronchoconstriction, the prostaglandin D2 receptor (gene: *PTGDR*) is thought to regulate, at least in part, levels of leukotriene C_4. In a study of 100 asthmatic children prescribed montelukast (5 mg/day), a modest effect of the *PTGDR*-4441T/C was observed.

Similarly, in a more extensive study of 169 SNPs in 26 candidate genes in asthma patients prescribed fluticasone, fluticasone propionate plus salmeterol or montelukast, a significant association between four SNPs in the cholinergic muscarinic receptor 2 gene (*CHRM2*) and montelukast response (change in FEV_1) over 16 weeks was observed [106]. *CHRM2* and *CHRM3* are thought to mediate airway tone via regulation of contractile/relaxation responses and targeting of *CHRM3* using antagonists; for example, tiotropium has shown clinical efficacy in multiple respiratory diseases. These data potentially suggest that the nature of the airway obstruction/tone determined by altered *CHRM2* expression and/or activity may influence responses to montelukast.

While the majority of studies have investigated the association between gene polymorphisms and acute responses to asthma therapy, a recent study by Mougey et al. identified that solute carrier organic anion transporter family, member 2B1 (*SLCO2B1*, alternative name: organic anion transported SB1; *OATP2B1*), was able to mediate montelukast permeability using a model cell system engineered to express the OATP2B1 protein [100]. Subsequently, the same group also showed that an alteration in the protein structure of OATP2B1 that occurs naturally—that is, Arg312Gln (rs12422149)—was associated with reduced morning plasma concentrations of montelukast following an evening dose during one month or six months of treatment in 80 asthma patients [100]. More specifically, GA (Arg/Gln) genotype carriers had ∼20% lower montelukast concentration than the GG (Arg) group at one month, and ∼30% lower concentration at six months, and of clinical relevance A allele carriers did not demonstrate benefit from montelukast using a symptom-based score [100].

The same researchers have now replicated these findings on montelukast absorption [109]; however, another study did not replicate this finding and failed to identify any effect of the Arg312Gln polymorphism on montelukast plasma concentrations, albeit in a different study design involving fewer subjects [101]. These studies highlight the potential importance of genetic factors influencing drug transporters that may be anticipated to affect pharmacokinetics and the pharmacodynamics. Similarly, several SNPs in another transporter protein, multidrug resistance protein 1 (*MRP1*) (alternative name: ATP binding cassette, subfamily C, member 1 (*ABCC1*))—for example, rs119774—have been associated with montelukast response (change in FEV_1 % predicted) [110] and zileuton response [103]. This association potentially confirms the pharmacogenetic significance of this *MRP1* SNP marker. This *SNP* is intronic, and the underlying functional mechanism, including which genetic variant explains these effects, remains to be resolved.

Overall, there has been good progress in the pharmacogenetics of leukotriene modifier therapy in asthma, with SNPs in multiple leukotriene-synthesising enzymes (e.g., *ALOX5*) showing robust association with LTA and LTSI responses. Multiple genes in the leukotriene synthesis pathway and/or receptors have polymorphisms that have been associated with asthma susceptibility, which implies that there may be a more leukotriene-driven asthma that is therefore more amenable to treatment with LTSIs and LTRAs. Of interest in this drug class, preliminary data suggest a significant contribution of SNPs in drug transporter genes, ultimately determining both the pharmacokinetics and the pharmacodynamics. These data set the scene for larger prospective studies to provide accurate effect sizes and determine clinical implications with greater confidence.

28.5.8 Corticosteroid Pharmacogenetics

Corticosteroids are an important pharmacogenetic target in asthma. Initial investigations into a pharmacogenetic approach for the corticosteroid element of asthma therapy focused on the glucocorticoid receptor gene (*GR*, alternative name: nuclear receptor subfamily 3, group C, member 1, *NR3C1*), which maps to chromosomal region 5q31—a region associated with multiple asthma phenotypes. As with other genes discussed in this chapter, several polymorphisms have been described in the *GR* gene that have functional consequences, such as a Val641Asp polymorphism that has been shown to influence the binding affinity for dexamethasone. However, similar to most polymorphisms shown to have potential pharmacogenetic effects, several of these polymorphisms are rare and their functional significance is questionable [84]. Despite this, the *GR* remains a tantalizing target for asthma pharmacogenetics, and it is surprising that there have not been further recent investigations into the role of GR polymorphisms in corticosteroid responses.

GR is not the only gene to be associated with genetic-based variance in corticosteroid efficacy in asthma treatment. A number of other genes have been implicated, including *CRHR1*, *TBX1*, *NK2R*, *STIP1*, *DUSP1*, and *FCER2*, and these are discussed below.

28.5.8.1 Corticotrophin-Releasing Hormone Receptor 1 Gene

SNPs in the corticotrophin-releasing hormone receptor 1 (*CRHR1*) gene are associated with response to inhaled corticosteroid treatment, based on end point change in FEV_1 following 8 weeks of treatment in three asthmatic cohorts using a candidate gene approach (131 SNPs in 14 genes, $n = 1117$ asthma subjects). For example, the intronic *CRHR1* SNP rs242941 exhibited genotype-specific changes in the percentage predicted change in FEV_1 in response to corticosteroid therapy in 470 adult asthma subjects [111]. This study was the first to show a pharmacogenetic effect for steroid efficacy in an asthmatic cohort, and it highlights the *CRHR1* gene to be at least one of a number of factors determining corticosteroid efficacy. *CRHR1* is thought to influence responses to exogenously administered corticosteroid through the regulation of endogenous levels of corticosteroid.

28.5.8.2 T Box 21 and Neurokinin 2 Receptor Genes

The T box 21 (*TBX21*) gene has been shown to be a predictor of improvement in bronchial hyper-responsiveness (BHR) (4 year change) post-corticosteroid treatment in children, based on a genotype-dependant variation of the nonsynonymous SNP rs2240017 (His33Gln), where the presence of the G allele gave greatest improvements [112]. However, the association of this gene remains under question because of its low allele frequency in Caucasians (MAF ∼0.04), which resulted in only a few subjects ($n = 5$) being available to contribute to this observation [112]. Additional data have given more confidence to the involvement of *TBX21* because the His33Gln polymorphism was significantly associated with improved asthma control in the presence of corticosteroids, although the alternative allele to that described was associated with greater control in 53 Korean asthma patients during 5–12 weeks treatment [112]. Interestingly in the same study, Ye et al. identified a novel gene, the Neurokinin 2 receptor (*NK2R*), which was associated with improved asthma control in the presence of corticosteroids; the G allele (Gly) of the *NK2R* SNP rs77038916G/A, (Gly231Glu) was associated with the greatest improvement. Neurokinin A induces bronchoconstriction and inflammation, therefore modulation of the neurokinin receptor may influence the magnitude of these responses.

28.5.8.3 Stress-Induced Phosphoprotein 1 Gene

Multiple SNPs in the stress-induced phosphoprotein 1 gene (*STIP1*) were associated with variable FEV_1 responses to treatment with the inhaled corticosteroid flunisolide in 382 asthma subjects [103]. This study investigated SNPs spanning eight candidate genes and identified *STIP1* SNPs rs4980524, rs6591838, and rs2236647 as SNPs affecting percent change in FEV_1 in response to flunisolide at 4 weeks and 8 weeks [103]. *STIP1* codes for an adaptor protein that coordinates functions with HSP70 and may be involved in formation of the glucocorticosteroid receptor heterocomplex.

28.5.8.4 Dual-Specificity Phosphatase 1 Gene

The dual-specificity phosphatase 1 (*DUSP1*) gene encodes a protein that has dual specificity for tyrosine and threonine and inactivates p38 mitogen-activated protein kinase (MAPK). Recently, several SNPs including a 5′ region SNP, rs881152, were associated with (1) bronchodilator responses and (2) asthma control in the presence of corticosteroid in a cohort of asthma patients (n = 430) [113]. The mechanism underlying these effects is unclear; however, it was suggested that corticosteroid induces *DUSP1* expression, which may be altered in carriers of the rs881152 5′-region SNP. This influences the ability of *DUSP1* to target the p38 MAPK signaling pathway.

28.5.8.5 Low-Affinity IgE Receptor Gene

In a cohort of 311 asthmatic children, the low-affinity IgE receptor gene (*FCER2*) has been highlighted as putatively involved in corticosteroid (budesonide) regulation of exacerbation by virtue of an association between the SNP T2206C (rs28364072, intronic) and the relative risk of severe exacerbations while taking budesonide [114]. Interestingly, the C allele was also associated with elevated IgE levels in the 311 asthma patients [114]. Importantly, this pharmacogenetic association has now been replicated in two other asthma cohorts (n = 386 and 939, respectively), with the 2206C allele being associated with increased hospital visits and uncontrolled asthma in patients receiving ICS [115]. It is not surprising that this polymorphic variation in an IgE receptor influences receptor expression, which alters IgE levels and is associated with more severe forms of asthma. In turn, these subjects cannot be adequately controlled by corticosteroid treatment.

The first GWAS in asthma, published in 2012, identified SNPs in the glucocorticoid-induced transcript 1 gene (*GLCC1*) as determinants of glucocorticoid response in 935 asthma subjects [116]. This study examined 534,290 SNPs using the Human Hap 550v3 Beadchip to identify 100 SNPs of interest in an initial cohort of 403 parent–child trios. The primary outcome was change in FEV_1. Of interest was a SNP in the *GLCC1* gene promoter region (rs37972C/T) that was associated with attenuation in FEV_1 improvement post-corticosteroid in three of four cohorts tested [116].

A subsequent GWAS identified the *T-gene* as a novel player in lung function response (FEV_1) to inhaled corticosteroids. Analyses were carried out in 418 Caucasian asthmatics and replicated in a secondary asthmatic population (n = 407), where 3 out of 47 successfully replicated SNPs were associated under the same genetic model in the same direction, including 2 of the top 4 SNPs ranked by p value. These SNPs (rs3127412C/T and rs6456042A/C) were in strong LD with a variant located in the *T-gene*, suggesting that this gene is a novel player in the pharmacogenetic regulation of corticosteroid efficacy. This was confirmed by a follow-up association, where the *T-gene* variant was associated with lung function response to inhaled corticosteroids in the initial GWAS analysis (an average 2–3-fold increase in FEV_1 response for homozygous wild-type subjects; rs3127412 TT; and rs6456042 CC) [117].

Not all recent developments have been due to GWAS. For example, a recent hypothesis-driven study implicated the cytochrome P450 3A enzymes in glucocorticosteroid efficacy through modulations in corticosteroid metabolism. In this study, Roberts et al. showed, through mRNA-driven studies confirmed via microsomes and recombinant enzymes, that CYP3A4 and CYP3A5 metabolize corticosteroids (budesonide) into inactive metabolites [118]. They suggested that differences in the expression or function of these enzymes in the lung and/or liver could influence corticosteroid efficacy in the treatment of asthma.

28.5.8.6 Combination Therapy (β_2-Adrenergic Receptor–Agonist/Corticosteroid) Pharmacogenetics

To date, the majority of pharmacogenetic studies have focused on one class of asthma therapy. However, several recent studies have investigated combination therapy, which more accurately reflects the clinical situation. We have summarized the main findings regarding the influence of *ADRB2* polymorphisms, including Arg16Gly, on combination therapies, particularly fluticasone/salmeterol and budesonide/formoterol in Section 28.5.5.1. However, it is clear that reports on associations between *ADRB2* Arg16 and increased exacerbation in asthma patients using regular salbutamol in addition to their maintenance combination are mixed, with some studies failing to report the association, potentially because of differences in study design.

More recently, additional genes have been investigated for SNP associations with patient responses to combination therapy. In a study of 81 asthma patients regularly using an ICS/LABA combination, it was shown that polymorphisms in the nitrous oxide synthase 3 (*NOS3*) gene involving an amino acid change (G894T, Glu298Asp), were associated with post-drug improvement in lung function (change in FEV_1 (% Pred)) [119]. *NOS3* codes for a key enzyme in the production of NO in the airways, and F_ENO has been shown to be elevated in asthma and is considered a marker of ongoing inflammation.

Importantly, earlier data identified that the Glu298Asp polymorphism is functional in *NOS3* and influences levels of F_ENO with the TT genotype associated with lower F_ENO levels. These data suggest that lower *NOS3* activity and potentially lower F_ENO levels identify patients more likely to respond to combination therapy; however, larger prospective trials are required to validate these findings because they are potentially counterintuitive if NO is a marker of inflammation and ICS target inflammation.

A recent 16-week study on the effect of genetic determinants (169 SNPs in 26 candidate genes) on asthma patient response to combined fluticasone propionate/salmeterol therapy identified three SNPs in the cholinergic muscarinic receptor 2 gene (*CHRM2*) that are associated with asthma ACQ scores, and found a single SNP, rs1461496 in heat shock 70kD protein 8 (*HSPA8*), to be associated with change in FEV_1 [106]. Of interest, *CHRM2* SNPs were also associated with responses (lung function) to montelukast (see earlier), suggesting that the "tone" of the airways as determined by muscarinic receptors is important for therapies that target airway contraction. *HSPA8* is involved in protein folding in the cell, but it is important to note that *HSPA8* SNPs previously were not associated with response to ICS [120].

28.6 DRUG DEVELOPMENT IN ASTHMA AND EMERGING TARGETS

Drug development in asthma has been slow to generate new classes, and the major advances in patient care have come from new compounds and/or indications for existing drug classes (e.g., improved duration of action for LABAs). One potential explanation for this limited progress is the design of and recruitment to Phase I and Phase II trials for the evaluation of new compounds. We have already outlined how initial trials of anti-IL5 demonstrated disappointing effects on lung function and asthma symptoms. Only after careful selection of patients, based on clinical and cellular parameters, were clinically relevant improvements observed. Researchers now understand that asthma is heterogeneous and, particularly for strategies targeting specific mediators, careful selection of patients is necessary to fully understand the therapeutic potential of new drugs.

Recent approaches to the treatment of asthma [121] include the following:

- Cytokine inhibition—for example, anti-IL5 (mepolizumab), anti-tumor necrosis factor alpha (TNFα) (etanercept, soluble TNFR), anti-Interleukin 13 (IL13) (lebrikizumab, tralokinumab)
- Chemokine receptor antagonists—for example, chemokine (CXC motif) receptor 2 (CXCR2) (SCH527123)
- Signaling pathway inhibitors with anti-inflammatory effects—for example, targeting p38 MAPK (SD-282), PDE4 (roflumilast)

The specific targeting of single mediators is unlikely to provide a therapeutic option for asthma in its broadest definition. However, targeting specific subpopulations has shown clinical efficacy. Very recently, a study investigating a human IL13-neutralizing monoclonal antibody (tralokinumab) used a carefully balanced population of atopic and nonatopic asthmatics, with exclusions for additional respiratory pathology, cigarette smoking ≥10 pack-years, recent infection, or treatment with immunosuppressive medication, in a randomized double-blind study [122]. Here, tralokinumab was associated with improved lung function based on an increase in FEV_1 and a decrease in daily β_2-adrenergic receptor–agonist use [122]. These studies overall demonstrate that careful selection of asthma subjects, in this case based on phenotype, is critical in asthma drug development.

As novel asthma genes are identified, it is paramount that their evaluation as potential drug targets, particularly in the context of Phase II trials, takes into account pharmacogenetic factors. Preliminary data support genetic testing in Phase II trials of newer compounds, as shown in a recent report on a IL4/IL13 dual antagonist (pitrikinra). Polymorphic variation in the target receptor for this antagonist (i.e., IL4Rα) significantly influenced outcomes in allergic asthma subjects [123]. Pitrikinra is a recombinant form of IL4 differing at two amino acid residues (i.e., a mutein, R121D/Y124D). This has been shown to reduce late-phase antigen responses (LAR) to inhaled antigen as defined by changes in lung function (FEV_1) over 4–10 hours post-antigen exposure.

In one study, subjects had been following a four-week period of twice daily active treatment of nebulized pitrikinra or placebo, with an increased LAR ratio correlating with a reduced FEV_1 [124]. Stratification of subjects (pitrikinra n = 15, placebo n = 14) into genotype groups for the nonsynonomous *IL4RA* SNP rs1801275 (Gln576Arg) identified Arg/Arg carriers as having an attenuated LAR ($P < 0.0001$) following pitrikinra treatment compared to Gln/Gln or Gln/Arg genotypes. Similarly, stratification based on rs1805011 (Glu400Ala) showed an attenuated LAR in the Glu/Ala group but not the Glu/Glu group [123]. Interestingly, the Arg576 variant (when in combination with Val75) was shown to be a risk factor for allergic asthma and to lead to enhanced IL4Rα signaling post IL4 stimulation [125].

These preliminary data suggest that selecting a subgroup of patients with a particular genotype when it is anticipated that the receptor/pathway may have a more dominant role in that individual's asthma is critical to interpreting Phase II clinical trials.

28.6.1 Screening, Stratification, Personalization

Approaches to further define asthma subphenotypes (e.g., using cluster analyses) have been useful and further confirm this heterogeneity in clinical presentation. While current medications have been extremely successful in asthma management (e.g., ICS), it is clear that individual responses to these medications and to therapies in development are heterogeneous, with large variation in clinical benefits and/or detrimental effects. The stratification of asthma treatment using clinical and/or genetic approaches therefore shows great potential for maximizing clinical outcomes and minimizing adverse effects, leading to improvement in the management of asthma.

There has been excellent progress in this area of research over the last five to ten years with the introduction of anti-IgE (omalizumab) therapy into mainstay asthma treatment, but only for those patients with clinical indications (i.e., severe allergic asthma) with an elevated serum IgE level. This stratification in prescribing for anti-IgE shows proof-of-concept by targeting this expensive biologic to those patients most likely to show clinical benefit. Similarly, there is accumulating evidence that stratification of patients in Phase II trials of newer therapies—for example, anti-IL5 (mepoluzimab) and anti-IL13 (lebrikizumab), again based on clinical and cellular patient profiles, is essential to adequately evaluate these therapies and will be important if these therapies are introduced into clinical practice.

Clear progress in asthma pharmacogenetics has come with completion of newer studies that interrogate multiple SNPs in large prospective cohorts. To date, predominantly candidate gene/pathway approaches have been used, but several genetic variants have now been identified with confidence (e.g., *ALOX5* and leukotriene modifier response; *FCER2* and glucocorticoid response). Importantly, these studies show independent replication (the gold standard in genetic studies, including specific SNPs and direction of effect).

Also importantly, several GWASs have now been completed that provide integration of common variation spanning the entire genome (typically testing 500,000 SNPs) identifying novel gene variants in a hypothesis-free approach (e.g., the collagen (*COL22A1*) locus and *SPATS2L* and response to albuterol/salbutamol). These more recent findings still need further replication and validation; however, they potentially pave the way for a "responder" and "nonresponder" profile based on multiple SNPs in multiple genes, all measured on the same platform. A key question is the relative contribution of common variation (measured on these platforms) and rare variation (not currently measured to these outcomes).

Another remaining question relates to the magnitude of changes in clinical measures between genotype groups in asthma patients (e.g., FEV_1). Data so far suggests that these differences can be considered clinically relevant; however, further work with larger populations is required to accurately estimate effect sizes. Similarly, the genetic variation identified to date accounts only for a modest proportion of the overall variation of the trait (e.g., a recent estimate suggests the hereditability of bronchodilator response at ~28.5%). It is likely that studies so far have been confounded by factors including SNP/haplotype analyses, gene–gene interactions, and the relative contribution of SNPs in different ethnic backgrounds. Therefore, while our knowledge has dramatically increased, there is a need for prospective trials of these current compounds that involve large numbers of subjects and hypothesis-free approaches including GWAS.

As with GWAS approaches to identifying disease susceptibility genes, "missing hereditability" will be an important issue in pharmacogenetics studies. This small proportion of hereditability currently assigned to identified loci (e.g., *CRHR1* < 3%) of corticosteroid response needs to be addressed because asthma medications cannot be successfully personalized unless we resolve a larger degree of genetic variability and/or environmental factors.

As with clinical/cellular approaches to stratifying Phase II trials of newer asthma therapies, it is clear that genetic stratification has a role to play. For example, the recent Phase II trials of the IL4/IL13 dual antagonist (pitrikinra) suggests patient selection is critical to evaluating the clinical efficacy of this compound, with specific IL4Rα genotypes identifying responder groups. More pharmacogenetic integration is required for Phase II trials, which represents a clear challenge for drug development pipelines.

As our understanding of the genetic basis of asthma increases other therapeutic targets for asthma will become apparent (e.g., IL33 receptor antagonists). It is clear that the genetic variation associated with asthma susceptibility in these genes will influence pathways targeting methods, as they essentially identify individuals in which the pathway may be particularly important. It is also likely that existing therapies may be influenced by the genetic factors that underlie asthma susceptibility, particularly therapies that have multiple targets (e.g., ICS).

The impact of our genetic knowledge on asthma management and prescribing practice is limited at this time, as there is a need to further define the complex genetic basis of responders and adverse effects in large prospective studies. However, preliminary "real-life" studies with related outcome measures (e.g., days from school) have provided proof of concept that genetic information can have important implications for disease management (e.g., informing second-line therapy in asthma using *ADRB2* Arg16) [98].

28.7 CONCLUSIONS AND FUTURE DIRECTIONS

There remain many unanswered questions regarding the potential of stratified approaches in the management of asthma. While progress has been made in identifying patient subgroups based on clinical measures, cell counts, or marker expression, more work is required using very large studies to extensively characterize asthma patients not only cross-sectionally, but also longitudinally over many years to identify the natural progression of the disease.

The field of genetics is rapidly progressing, with great technological developments moving at a dramatic pace, e.g., GWAS comprising a million common SNPs or several million rare SNPs, gene expression studies, and whole-exome or targeted resequencing. These approaches will allow the interrogation of genetic variation effect in patient subgroups. It is likely that a drug-specific genetic profile involving several genes (e.g., receptors, signaling mediators, transcription factors) will be a step toward personalized medicine in asthma, with associated benefits including avoidance of adverse side effects and adequate control of the disease. We anticipate that through these approaches many novel common variants/genes will be identified as underlying responses to current asthma medication. As the cost of genetic analyses continues to fall, the potential of a relatively simple test that will have an impact on asthma management costs becomes a real possibility.

If genetic or clinical information is going to influence clinical practice, there must be a simple, reproducible diagnostic test based on SNP combinations or simple clinical measures, with adequately high specificity and sensitivity that distinguishes responders and nonresponders. However, the introduction of routine genetic and/or clinical testing into practice will require a clear demonstration of health benefits and cost-effectiveness that outperforms the current stepwise approach to asthma management.

ACKNOWLEDGMENTS

Research in the authors' laboratory is funded by Asthma UK, the Medical Research Council, and the British Medical Association. We thank Dr Emily Hodge for originally generating Figure 28.2.

REFERENCES

[1] A plea to abandon asthma as a disease concept. Lancet 2006; 368(9537):705.

[2] Holgate ST. A brief history of asthma and its mechanisms to modern concepts of disease pathogenesis. Allergy Asthma Immunol Res 2010;2(3):165–71 [Epub 2010/07/02].

[3] National Heart Lung, and Blood Institute. Guidelines for the diagnosis and management of asthma. National asthma education and prevention program: expert panel report III. NIH Publication No. 08-4051; 2007.

[4] British Thoracic Society and Scottish Intercollegiate Guidelines Network: British Guideline on the Management of Asthma. A national clinical guideline; 2012 [Revision].

[5] Bahadori K, Doyle-Waters MM, Marra C, Lynd L, Alasaly K, Swiston J, et al. Economic burden of asthma: a systematic review. BMC Pulm Med 2009;9:24 [Epub 2009/05/21].

[6] van Ganse E, Antonicelli L, Zhang Q, Laforest L, Yin DD, Nocea G, et al. Asthma-related resource use and cost by GINA classification of severity in three European countries. Respir Med 2006;100(1):140–7 [Epub 2005/12/13].

[7] Price D, Bosnic-Anticevich S, Briggs A, Chrystyn H, Rand C, Scheuch G, et al. Inhaler competence in asthma: common errors, barriers to use and recommended solutions. Respir Med 2013;107(1):37–46.

[8] Suissa S, Ernst P, Benayoun S, Baltzan M, Cai B. Low-dose inhaled corticosteroids and the prevention of death from asthma. N Engl J Med 2000;343(5):332–6.

[9] Bousquet J. Global initiative for asthma (GINA) and its objectives. Clin Exp Allergy 2000;30(Suppl. 1):2–5.

[10] Reddel HK, Taylor DR, Bateman ED, Boulet LP, Boushey HA, Busse WW, et al. An official American thoracic society/European respiratory society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. Am J Respir Crit Care Med 2009;180(1):59–99.

[11] Reddel H, Ware S, Marks G, Salome C, Jenkins C, Woolcock A. Differences between asthma exacerbations and poor asthma control. Lancet 1999;353(9150):364–9.

[12] Hoskins G, McCowan C, Neville RG, Thomas GE, Smith B, Silverman S. Risk factors and costs associated with an asthma attack. Thorax 2000;55(1):19–24.

[13] Wark PA, Johnston SL, Moric I, Simpson JL, Hensley MJ, Gibson PG. Neutrophil degranulation and cell lysis is associated with clinical severity in virus-induced asthma. Eur Respir J 2002;19(1):68–75 [Epub 2002/02/21].

[14] Papadopoulos NG, Christodoulou I, Rohde G, Agache I, Almqvist C, Bruno A, et al. Viruses and bacteria in acute asthma exacerbations—a GA2LEN-DARE* systematic review. Allergy 2011;66(4):458–68.

[15] Talbot TR, Hartert TV, Mitchel E, Halasa NB, Arbogast PG, Poehling KA, et al. Asthma as a risk factor for invasive pneumococcal disease. N Engl J Med 2005;352(20):2082–90.

[16] Biscione GL, Corne J, Chauhan AJ, Johnston SL. Increased frequency of detection of *Chlamydophila pneumoniae* in asthma. Eur Respir J 2004;24(5):745–9.

[17] Biscardi S, Lorrot M, Marc E, Moulin F, Boutonnat-Faucher B, Heilbronner C, et al. *Mycoplasma pneumoniae* and asthma in children. Clin Infect Dis 2004;38(10):1341–6.

[18] Johnston SL, Blasi F, Black PN, Martin RJ, Farrell DJ, Nieman RB, et al. The effect of telithromycin in acute exacerbations of asthma. N Engl J Med 2006;354(15):1589–600.

[19] Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. Disordered microbial communities in asthmatic airways. PLoS ONE 2010;5(1):e8578.

[20] Sykes A, Johnston SL. Etiology of asthma exacerbations. J Allergy Clin Immunol 2008;122(4):685–8.

[21] Marks GB, Colquhoun JR, Girgis ST, Koski MH, Treloar AB, Hansen P, et al. Thunderstorm outflows preceding epidemics of asthma during spring and summer. Thorax 2001;56(6):468–71 [Epub 2001/05/22].

[22] Denning DW, O'Driscoll BR, Hogaboam CM, Bowyer P, Niven RM. The link between fungi and severe asthma: a summary of the evidence. Eur Respir J 2006;27(3):615–26 [Epub 2006/03/02].

[23] O'Hollaren MT, Yunginger JW, Offord KP, Somers MJ, O'Connell EJ, Ballard DJ, et al. Exposure to an aeroallergen as a possible precipitating factor in respiratory arrest in young patients with asthma. N Engl J Med 1991;324(6):359–63 [Epub 1991/02/07].

[24] Singh AM, Busse WW. Asthma exacerbations 2: aetiology. Thorax 2006;61(9):809–16.

[25] Denning DW, O'Driscoll BR, Powell G, Chew F, Atherton GT, Vyas A, et al. Randomized controlled trial of oral antifungal treatment for severe asthma with fungal sensitization: the fungal asthma sensitization trial (FAST) study. Am J Respir Crit Care Med 2009;179(1):11–8.

[26] Shaw D, Green R, Berry M, Mellor S, Hargadon B, Shelley M, et al. A cross-sectional study of patterns of airway dysfunction, symptoms and morbidity in primary care asthma. Prim Care Respir J 2012;21(3):283–7 [Epub 2012/07/13].

[27] Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. Eur Respir J 2005;26(5):948–68 [Epub 2005/11/03].

[28] Lang A, Carlsen KH, Haaland G, Devulapalli CS, Munthe-Kaas M, Mowinckel P, et al. Severe asthma in childhood: assessed in 10 year olds in a birth cohort study. Allergy 2008;63(8):1054–60 [Epub 2008/08/12].

[29] Strippoli M-PF, Spycher BD, Pescatore AM, Beardsmore CS, Silverman M, Kuehni CE. Exclusive viral wheeze and allergic wheeze: evidence for discrete phenotypes. Eur Respir J 2011;38(2):472–4.

[30] Bush A, Fleming L. Phenotypes of refractory/severe asthma. Paediatr Respir Rev 2011;12(3):177–81 [Epub 2011/07/05].

[31] Kermode JA, Brown NJ, Hardaker KM, Farah CS, Berend N, King GG, et al. The effect of airway remodelling on airway hyperresponsiveness in asthma. Respir Med 2011;105(12):1798–804 [Epub 2011/08/09].

[32] Hodge E, Sayers I. Allergy. In: eLS. Chichester: John Wiley & Sons Ltd; 2013.

[33] Brusasco V, Crimi E, Pellegrino R. Airway hyperresponsiveness in asthma: not just a matter of airway inflammation. Thorax 1998;53(11):992–8 [Epub 1999/04/08].

[34] Barnes PJ. Is asthma a nervous disease? The Parker B. Francis lectureship. Chest 1995;107(Suppl. 3):119S–25S [Epub 1995/03/01].

[35] Laitinen LA, Laitinen A, Haahtela T. Airway mucosal inflammation even in patients with newly diagnosed asthma. Am Rev Respir Dis 1993;147(3):697–704 [Epub 1993/03/01].

[36] O'Donnell RA, Davies DE, Holgate ST. Airway remodelling. In: Barnes P, Drazen J, Rennard S, Thomson N, editors. Asthma and COPD. London: Academic Press; 2002.

[37] Jeffery PK. Structural and inflammatory changes in COPD: a comparison with asthma. Thorax 1998;53(2):129–36 [Epub 1998/06/13].

[38] Grainge CL, Lau LC, Ward JA, Dulay V, Lahiff G, Wilson S, et al. Effect of bronchoconstriction on airway remodeling in asthma. N Engl J Med 2011;364(21):2006–15 [Epub 2011/05/27].

[39] Cooke RA, van der Veer A. Human sensitisation. J Immunol 1916;1:201–305.

[40] Gerrard JW, Vickers P, Gerrard CD. The familial incidence of allergic disease. Ann Allergy 1976;36(1):10–15 [Epub 1976/01/01].

[41] Skadhauge LR, Christensen K, Kyvik KO, Sigsgaard T. Genetic and environmental influence on asthma: a population-based study of 11,688 Danish twin pairs. Eur Respir J 1999;13(1):8–14 [Epub 2000/06/03].

[42] Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. Immunol Rev 2011;242(1):10–30 [Epub 2011/06/21].

[43] Akhabir L, Sandford AJ. Genome-wide association studies for discovery of genes involved in asthma. Respirology 2011;16(3):396–406 [Epub 2011/02/01].

[44] van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. Nature 2002;418(6896):426–30 [Epub 2002/07/12].

[45] Barton SJ, Koppelman GH, Vonk JM, Browning CA, Nolte IM, Stewart CE, et al. PLAUR polymorphisms are associated with asthma, PLAUR levels, and lung function decline. J Allergy Clin Immunol 2009;123(6):1391–400 e17 [Epub 2009/05/16].

[46] Koppelman GH, Sayers I. Evidence of a genetic contribution to lung function decline in asthma. J Allergy Clin Immunol 2011;128(3):479–84 [Epub 2011/07/15].

[47] Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. Genes Immun 2006;7(2):95–100 [Epub 2006/01/06].

[48] Undarmaa S, Mashimo Y, Hattori S, Shimojo N, Fujita K, Miyatake A, et al. Replication of genetic association studies in asthma and related phenotypes. J Hum Genet 2010;55(6):342–9 [Epub 2010/04/17].

[49] Wills-Karp M. Interleukin-13 in asthma pathogenesis. Immunol Rev 2004;202:175–90.

[50] Humbert M, Durham SR, Kimmitt P, Powell N, Assoufi B, Pfister R, et al. Elevated expression of messenger ribonucleic acid encoding IL-13 in the bronchial mucosa of atopic and nonatopic subjects with asthma. J Allergy Clin Immunol 1997;99(5):657–65.

[51] Wills-Karp M, Luyimbazi J, Xu X, Schofield B, Neben TY, Karp CL, et al. Interleukin-13: central mediator of allergic asthma. Science 1998;282(5397):2258–61.

[52] Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 2007;448(7152):470–3 [Epub 2007/07/06].

[53] Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association

study of asthma. N Engl J Med 2010;363(13):1211–21 [Epub 2010/09/24].

[54] Wan YI, Shrine NR, Soler Artigas M, Wain LV, Blakey JD, Moffatt MF, et al. Genome-wide association study to identify genetic determinants of severe asthma. Thorax 2012;67(9):762–8 [Epub 2012/05/09].

[55] Kearley J, Buckland KF, Mathie SA, Lloyd CM. Resolution of allergic inflammation and airway hyperreactivity is dependent upon disruption of the T1/ST2-IL-33 pathway. Am J Respir Crit Care Med 2009;179(9):772–81 [Epub 2009/01/31].

[56] Hammad H, Chieppa M, Perros F, Willart MA, Germain RN, Lambrecht BN. House dust mite allergen induces asthma via Toll-like receptor 4 triggering of airway structural cells. Nat Med 2009;15(4):410–6 [Epub 2009/03/31].

[57] Oshikawa K, Kuroiwa K, Tago K, Iwahana H, Yanagisawa K, Ohno S, et al. Elevated soluble ST2 protein levels in sera of patients with asthma with an acute exacerbation. Am J Respir Crit Care Med 2001;164(2):277–81 [Epub 2001/07/21].

[58] Raissy HH, Kelly HW, Harkins M, Szefler SJ. Inhaled corticosteroids in lung diseases. Am J Respir Crit Care Med 2013;187(8):798–803.

[59] Rackemann FM. Intrinsic asthma. J Nerv Ment Dis 1947;106(4):497 [Epub 1947/10/01].

[60] Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, et al. Cluster analysis and clinical asthma phenotypes. Am J Respir Crit Care Med 2008;178(3):218–24.

[61] Siroux V, Basagana X, Boudier A, Pin I, Garcia-Aymerich J, Vesin A, et al. Identifying adult asthma phenotypes using a clustering approach. Eur Respir J 2011;38(2):310–7 [Epub 2011/01/15].

[62] Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, et al. Identification of asthma phenotypes using cluster analysis in the severe asthma research program. Am J Respir Crit Care Med 2010;181(4):315–23.

[63] Vijverberg SJ, Koenderman L, van Erp FC, van der Ent CK, Postma DS, Brinkman P, et al. Inflammatory phenotypes underlying uncontrolled childhood asthma despite inhaled corticosteroid treatment: rationale and design of the PACMAN2 study. BMC Pediatr 2013;13(1):94 [Epub 2013/06/19].

[64] Humbert M, Menz G, Ying S, Corrigan CJ, Robinson DS, Durham SR, et al. The immunopathology of extrinsic (atopic) and intrinsic (non-atopic) asthma: more similarities than differences. Immunol Today 1999;20(11):528–33.

[65] Malmstrom K, Rodriguez-Gomez G, Guerra J, Villaran C, Pineiro A, Wei LX, et al. Oral montelukast, inhaled beclomethasone, and placebo for chronic asthma. A randomized, controlled trial. Montelukast/Beclomethasone Study Group. Ann Intern Med 1999;130(6):487–95 [Epub 1999/03/13].

[66] Gruffydd-Jones K. Measuring pulmonary function in practice. Practitioner 2002;246(1635):445–9 [Epub 2002/06/21].

[67] Drazen JM, Silverman EK, Lee TH. Heterogeneity of therapeutic responses in asthma. Br Med Bull 2000;56(4):1054–70 [Epub 2001/05/22].

[68] Fahy JV, Fleming HE, Wong HH, Liu JT, Su JQ, Reimann J, et al. The effect of an anti-IgE monoclonal antibody on the early- and late-phase responses to allergen inhalation in asthmatic subjects. Am J Respir Crit Care Med 1997;155(6):1828–34 [Epub 1997/06/01].

[69] Hanania NA, Wenzel S, Rosen K, Hsieh HJ, Mosesova S, Choy DF, et al. Exploring the effects of omalizumab in allergic asthma. Am J Respir Crit Care Med 2013;187(8):804–11 [Epub 2013/03/09].

[70] Berry MA, Morgan A, Shaw DE, Parker D, Green RH, Brightling CE, et al. Pathological features and inhaled corticosteroid response of eosinophilic and non-eosinophilic asthma. Thorax 2007;62(12):1043–9.

[71] Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, et al. T-helper Type 2–driven inflammation defines major subphenotypes of asthma. Am J Respir Crit Care Med 2009;180(5):388–95.

[72] Szefler SJ, Phillips BR, Martinez FD, Chinchilli VM, Lemanske RF, Strunk RC, et al. Characterization of within-subject responses to fluticasone and montelukast in childhood asthma. J Allergy Clin Immunol 2005;115(2):233–42.

[73] Corren J, Lemanske RF, Hanania NA, Korenblat PE, Parsey MV, Arron JR, et al. Lebrikizumab treatment in adults with asthma. N Engl J Med 2011;365(12):1088–98.

[74] Leckie MJ, ten Brinke A, Khan J, Diamant Z, O'Connor BJ, Walls CM, et al. Effects of an interleukin-5 blocking monoclonal antibody on eosinophils, airway hyper-responsiveness, and the late asthmatic response. Lancet 2000;356(9248):2144–8.

[75] Haldar P, Brightling CE, Hargadon B, Gupta S, Monteiro W, Sousa A, et al. Mepolizumab and exacerbations of refractory eosinophilic asthma. N Engl J Med 2009;360(10):973–84.

[76] Tattersfield AE, Postma DS, Barnes PJ, Svensson K, Bauer CA, O'Byrne PM, et al. Exacerbations of asthma: a descriptive study of 425 severe exacerbations. The FACET international study group. Am J Respir Crit Care Med 1999;160(2):594–9.

[77] Rosi E, Ronchi MC, Grazzini M, Duranti R, Scano G. Sputum analysis, bronchial hyperresponsiveness, and airway function in asthma: results of a factor analysis. J Allergy Clin Immunol 1999;103(2 Pt 1):232–7.

[78] Birring SS, Berry M, Brightling CE, Pavord ID. Eosinophilic bronchitis: clinical features, management and pathogenesis. Am J Respir Med 2003;2(2):169–73.

[79] Hunter CJ, Brightling CE, Woltmann G, Wardlaw AJ, Pavord ID. A comparison of the validity of different diagnostic tests in adults with asthma. Chest 2002;121(4):1051–7.

[80] Sont JK, Willems LN, Bel EH, van Krieken JH, Vandenbroucke JP, Sterk PJ. Clinical control and histopathologic outcome of asthma when using airway hyperresponsiveness as an additional guide to long-term treatment. The AMPUL study group. Am J Respir Crit Care Med 1999;159(4 Pt 1):1043–51.

[81] Green RH, Brightling CE, McKenna S, Hargadon B, Parker D, Bradding P, et al. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. Lancet 2002;360(9347):1715–21.

[82] Green RH, Brightling CE, Woltmann G, Parker D, Wardlaw AJ, Pavord ID. Analysis of induced sputum in adults with asthma: identification of subgroup with isolated sputum neutrophilia and poor response to inhaled corticosteroids. Thorax 2002;57(10):875–9.

[83] Jayaram L, Pizzichini MM, Cook RJ, Boulet LP, Lemiere C, Pizzichini E, et al. Determining asthma treatment by monitoring sputum cell counts: effect on exacerbations. Eur Respir J 2006;27(3):483–94.

[84] Portelli M, Sayers I. Genetic basis for personalized medicine in asthma. Expert Rev Respir Med 2012;6(2):223–36 [Epub 2012/03/30].

[85] Reihsaus E, Innis M, MacIntyre N, Liggett SB. Mutations in the gene encoding for the beta 2-adrenergic receptor in normal and asthmatic subjects. Am J Respir Cell Mol Biol 1993;8(3):334–9 [Epub 1993/03/01].

[86] Lee MY, Cheng SN, Chen SJ, Huang HL, Wang CC, Fan HC. Polymorphisms of the beta2-adrenergic receptor correlated to nocturnal asthma and the response of terbutaline nebulizer. Pediatr Neonatol 2011;52(1):18–23 [Epub 2011/03/10].

[87] Basu K, Palmer CN, Tavendale R, Lipworth BJ, Mukhopadhyay S. Adrenergic beta(2)-receptor genotype predisposes to exacerbations in steroid-treated asthmatic patients taking frequent albuterol or salmeterol. J Allergy Clin Immunol 2009;124(6):1188–94 e3 [Epub 2009/10/06].

[88] Bleecker ER, Postma DS, Lawrance RM, Meyers DA, Ambrose HJ, Goldman M. Effect of ADRB2 polymorphisms on response to longacting beta2-agonist therapy: a pharmacogenetic analysis of two randomised studies. Lancet 2007;370(9605):2118–25 [Epub 2007/12/25].

[89] Panebra A, Wang WC, Malone MM, Pitter DR, Weiss ST, Hawkins GA, et al. Common ADRB2 haplotypes derived from 26 polymorphic sites direct beta2-adrenergic receptor expression and regulation phenotypes. PLoS ONE 2010;5(7):e11819 [Epub 2010/08/06].

[90] Litonjua AA, Lasky-Su J, Schneiter K, Tantisira KG, Lazarus R, Klanderman B, et al. ARG1 is a novel bronchodilator response gene: screening and replication in four asthma cohorts. Am J Respir Crit Care Med 2008;178(7):688–94 [Epub 2008/07/12].

[91] Vonk JM, Postma DS, Maarsingh H, Bruinenberg M, Koppelman GH, Meurs H. Arginase 1 and arginase 2 variations associate with asthma, asthma severity and beta2 agonist and steroid response. Pharmacogenet Genomics 2010;20(3):179–86 [Epub 2010/02/04].

[92] Duan QL, Gaume BR, Hawkins GA, Himes BE, Bleecker ER, Klanderman B, et al. Regulatory haplotypes in ARG1 are associated with altered bronchodilator response. Am J Respir Crit Care Med 2011;183(4):449–54 [Epub 2010/09/21].

[93] Moore PE, Ryckman KK, Williams SM, Patel N, Summar ML, Sheller JR. Genetic variants of GSNOR and ADRB2 influence response to albuterol in African-American children with severe asthma. Pediatr Pulmonol 2009;44(7):649–54 [Epub 2009/06/11].

[94] Choudhry S, Que LG, Yang Z, Liu L, Eng C, Kim SO, et al. GSNO reductase and beta2-adrenergic receptor gene-gene interaction: bronchodilator responsiveness to albuterol. Pharmacogenet Genomics 2010;20(6):351–8 [Epub 2010/03/26].

[95] Gaston B, Reilly J, Drazen JM, Fackler J, Ramdev P, Arnelle D, et al. Endogenous nitrogen oxides and bronchodilator S-nitrosothiols in human airways. Proc Natl Acad Sci USA 1993;90(23):10957–61 [Epub 1993/12/01].

[96] Whalen EJ, Foster MW, Matsumoto A, Ozawa K, Violin JD, Que LG, et al. Regulation of beta-adrenergic receptor signaling by S-nitrosylation of G-protein-coupled receptor kinase 2. Cell 2007;129(3):511–22 [Epub 2007/05/08].

[97] Duan QL, Lasky-Su J, Himes BE, Qiu W, Litonjua AA, Damask A, et al. A genome-wide association study of bronchodilator response in asthmatics. Pharmacogenomics J 2014;14(1):41–7 [Epub 2013/03/20].

[98] Lipworth BJ, Basu K, Donald HP, Tavendale R, Macgregor DF, Ogston SA, et al. Tailored second-line therapy in asthmatic children with the Arg(16) genotype. Clin Sci (Lond) 2013;124(8):521–8 [Epub 2012/11/07].

[99] Duroudier NP, Tulah AS, Sayers I. Leukotriene pathway genetics and pharmacogenetics in allergy. Allergy 2009;64(6):823–39 [Epub 2009/05/07].

[100] Mougey EB, Feng H, Castro M, Irvin CG, Lima JJ. Absorption of montelukast is transporter mediated: a common variant of OATP2B1 is associated with reduced plasma concentrations and poor response. Pharmacogenet Genomics 2009;19(2):129–38 [Epub 2009/01/20].

[101] Tapaninen T, Karonen T, Backman JT, Neuvonen PJ, Niemi M. SLCO2B1 c.935G>A single nucleotide polymorphism has no effect on the pharmacokinetics of montelukast and aliskiren. Pharmacogenet Genomics 2013;23(1):19–24 [Epub 2012/11/16].

[102] Drazen JM, Yandava CN, Dube L, Szczerback N, Hippensteel R, Pillari A, et al. Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. Nat Genet 1999;22(2):168–70.

[103] Hawkins GA, Lazarus R, Smith RS, Tantisira KG, Meyers DA, Peters SP, et al. The glucocorticoid receptor heterocomplex gene STIP1 is associated with improved lung function in asthmatic subjects treated with inhaled corticosteroids. J Allergy Clin Immunol 2009;123(6):1376–83. e7 [Epub 2009/03/04].

[104] Lima JJ, Zhang S, Grant A, Shao L, Tantisira KG, Allayee H, et al. Influence of leukotriene pathway polymorphisms on response to montelukast in asthma. Am J Respir Crit Care Med 2006;173(4):379–85 [Epub 2005/11/19].

[105] Klotsman M, York TP, Pillai SG, Vargas-Irwin C, Sharma SS, van den Oord EJ, et al. Pharmacogenetics of the 5-lipoxygenase biosynthetic pathway and variable clinical response to montelukast. Pharmacogenet Genomics 2007;17(3):189–96 [Epub 2007/04/27].

[106] Mougey EB, Chen C, Tantisira KG, Blake KV, Peters SP, Wise RA, et al. Pharmacogenetics of asthma controller treatment. Pharmacogenomics J 2013;13(3):242–50 [Epub 2012/03/01].

[107] Sampson AP, Siddiqui S, Buchanan D, Howarth PH, Holgate ST, Holloway JW, et al. Variant LTC(4) synthase allele modifies cysteinyl leukotriene synthesis in eosinophils and predicts clinical response to zafirlukast. Thorax 2000;55(Suppl. 2):S28–31 [Epub 2000/09/19].

[108] York TP, Vargas-Irwin C, Anderson WH, van den Oord EJ. Asthma pharmacogenetic study using finite mixture models to handle drug-response heterogeneity. Pharmacogenomics 2009;10(5):753–67 [Epub 2009/05/20].

[109] Mougey EB, Lang JE, Wen X, Lima JJ. Effect of citrus juice and SLCO2B1 genotype on the pharmacokinetics of montelukast. J Clin Pharmacol 2011;51(5):751–60 [Epub 2010/10/27].

[110] Lima JJ. Treatment heterogeneity in asthma: genetics of response to leukotriene modifiers. Mol Diagn Ther 2007;11(2):97–104 [Epub 2007/04/03].

[111] Tantisira KG, Lake S, Silverman ES, Palmer LJ, Lazarus R, Silverman EK, et al. Corticosteroid pharmacogenetics: association of sequence variants in CRHR1 with improved lung function in asthmatics treated with inhaled corticosteroids. Hum Mol Genet 2004;13(13):1353–9.

[112] Ye YM, Lee HY, Kim SH, Jee YK, Lee SK, Lee SH, et al. Pharmacogenetic study of the effects of NK2R G231E G>A and TBX21 H33Q C>G polymorphisms on asthma control with inhaled corticosteroid treatment. J Clin Pharm Ther 2009;34(6):693–701 [Epub 2010/02/24].

[113] Jin Y, Hu D, Peterson EL, Eng C, Levin AM, Wells K, et al. Dual-specificity phosphatase 1 as a pharmacogenetic modifier of inhaled steroid response among asthmatic patients. J Allergy Clin Immunol 2010;126(3):618–25 e1–2 [Epub 2010/08/03].

[114] Tantisira KG, Silverman ES, Mariani TJ, Xu J, Richter BG, Klanderman BJ, et al. FCER2: a pharmacogenetic basis for severe exacerbations in children with asthma. J Allergy Clin Immunol 2007;120(6):1285–91 [Epub 2007/11/06].

[115] Koster ES, Maitland-Vander Zee AH, Tavendale R, Mukhopadhyay S, Vijverberg SJ, Raaijmakers JA, et al. FCER2 T2206C variant associated with chronic symptoms and exacerbations in steroid-treated asthmatic children. Allergy 2011;66(12):1546–52. [Epub 2011/10/01]

[116] Tantisira KG, Lasky-Su J, Harada M, Murphy A, Litonjua AA, Himes BE, et al. Genomewide association between GLCCI1 and response to glucocorticoid therapy in asthma. N Engl J Med 2011;365(13):1173–83.

[117] Tantisira KG, Damask A, Szefler SJ, Schuemann B, Markezich A, Su J, et al. Genome-wide association identifies the T gene as a novel asthma pharmacogenetic locus. Am J Respir Crit Care Med 2012;185(12):1286–91 [Epub 2012/04/28].

[118] Roberts JK, Moore CD, Ward RM, Yost GS, Reilly CA. Metabolism of beclomethasone dipropionate by cytochrome P450 3A enzymes. J Pharmacol Exp Ther 2013;345(2):308–16 [Epub 2013/03/21].

[119] Iordanidou M, Paraskakis E, Tavridou A, Paschou P, Chatzimichael A, Manolopoulos VG. G894T polymorphism of eNOS gene is a predictor of response to combination of inhaled corticosteroids with long-lasting beta2-agonists in asthmatic children. Pharmacogenomics 2012;13(12):1363–72 [Epub 2012/09/13].

[120] Hawkins GA, Peters SP. Pharmacogenetics of asthma. Methods Mol Biol 2008;448:359–78 [Epub 2008/03/29].

[121] Barnes KC. Ancestry, ancestry-informative markers, asthma, and the quest for personalized medicine. J Allergy Clin Immunol 2010;126(6):1139–40 [Epub 2010/12/08].

[122] Piper E, Brightling C, Niven R, Oh C, Faggioni R, Poon K, et al. A phase II placebo-controlled study of tralokinumab in moderate-to-severe asthma. Eur Respir J 2013;41(2):330–8 [Epub 2012/06/30].

[123] Slager RE, Hawkins GA, Ampleford EJ, Bowden A, Stevens LE, Morton MT, et al. IL-4 receptor alpha polymorphisms are predictors of a pharmacogenetic response to a novel IL-4/IL-13 antagonist. J Allergy Clin Immunol 2010;126(4):875–8 [Epub 2010/10/06].

[124] Wenzel S, Wilbraham D, Fuller R, Getz EB, Longphre M. Effect of an interleukin-4 variant on late phase asthmatic response to allergen challenge in asthmatic patients: results of two phase 2a studies. Lancet 2007;370(9596):1422–31 [Epub 2007/10/24].

[125] Risma KA, Wang N, Andrews RP, Cunningham CM, Ericksen MB, Bernstein JA, et al. V75R576 IL-4 receptor alpha is associated with allergic asthma and enhanced IL-4 receptor function. J Immunol 2002;169(3):1604–10 [Epub 2002/07/23].

This page is intentionally left blank

Chapter 29

Rheumatoid Arthritis

Jisna R Paul, Deepali Sen, and Prabha Ranganathan
Division of Rheumatology, Department of Medicine, Washington University School of Medicine, St. Louis, USA

Chapter Outline

29.1 Overview 653
29.1.1 Epidemiology 653
29.1.2 Immunologic Features 653
29.1.3 Pathogenesis 654
29.1.4 Diagnosis 654
29.1.5 Treatment 654
29.1.6 Pharmacogenetics and Stratified Medicines in RA 655
29.2 Pharmacogenetics of Nonbiologic DMARDs 655
29.2.1 Methotrexate 655
29.2.2 Pharmacogenetics of Leflunomide 661
29.2.3 Pharmacogenetics of Azathioprine 663
29.2.4 Sulfasalazine 665
29.2.5 Pharmacogenetics of Hydroxychloroquine 666
29.3 Pharmacogenetics of Biologic DMARDS 668
29.3.1 Tumor Necrosis Factor Antagonists 668
29.3.2 Rituximab 673
29.3.3 Tocilizumab 674
29.4 Conclusions and Future Directions 674
References 674

29.1 OVERVIEW

Rheumatoid arthritis (RA) is an inflammatory symmetric polyarthritis that left untreated can lead to erosions, joint space loss, and permanent destruction of the affected joints. RA is a chronic, disabling disease and can be associated with systemic manifestations.

29.1.1 Epidemiology

RA affects approximately 1% of the population worldwide and leads to significant morbidity and mortality. It increases in incidence with age and affects women about three times more often than men [1]. However, after the age of 60, RA affects both genders equally.

29.1.2 Immunologic Features

The genetic basis of RA remains incompletely understood, although an association with major histocompatibility complex (MHC) class II molecules is well known. Specifically, individuals with the "shared epitope," a unique sequence of amino acids in the hypervariable region of the HLA-DR β chain, are at increased risk of RA than are those without it. *HLA-DRB1* alleles that code the shared epitope (SE)—a 5 amino acid–sequence motif in residues 70–74 of the HLA-DRβ chain—confer a strong genetic risk for RA [2,3]. Although the exact mechanisms for this association remain unclear, glutamine or arginine in position 70 on the DRβ chain increases RA risk, while aspartic acid in that position confers protection [4]. Up to 20% of RA patients are SE-negative [2]. Noninherited maternal HLA antigens (NIMA) may confer RA risk in some individuals [5,6].

Many environmental factors, in particular cigarette smoking and periodontal disease, have been identified as risk factors for RA [1]. This risk is further increased in individuals who also have the shared epitope [7]. Supporting the role of cigarette smoking in the development of RA is some evidence showing that smoking cessation may help prevent its development [8].

Deimination of arginine to citrulline by peptidyl arginine deiminase (PADI) is a well-described post-translational protein modification in RA and has been implicated in RA's pathophysiology. Various citrullinated proteins have been demonstrated to be in the rheumatoid joint, including fibrinogen, collagen, and fibronectin. The process of citrullination involves conversion of arginine to citrulline by the enzymes peptidylarginine deiminases (PADIs). Of the four isoforms, PADI2 and PADI4 are most abundant in

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00029-3
© 2014 Elserier Inc. All rights reserved.

the inflamed synovium. PADI4 polymorphisms may confer increased susceptibility to RA in Asians and Europeans [9]. In RA, increased citrullination occurs in the inflamed synovium, and antibodies directed against these citrullinated proteins (anticyclic citrullinated peptide [anti-CCP] antibodies [ACPA]) are produced by resident B cells. ACPA may be an important effector mechanism in RA disease pathogenesis [2].

29.1.3 Pathogenesis

Although the etiology of RA is still unknown, there have been major strides in understanding how the inflammatory process leads to joint destruction. Several processes contribute to joint destruction in RA, including synovial immune complex deposition, neutrophil infiltration, angiogenesis, and T-cell activation [10]. T-cell activation triggered by an unknown antigen leads to joint inflammation and destruction. Synovial membrane hyperplasia and increased vascularity lead to pannus formation. The pannus shows leukocytic infiltration, increased expression of adhesion molecules, proteolytic enzymes, and cytokines, including tumor necrosis factor (TNF)-α, and interleukins (IL)-1, 6, and 17 [11]. These cytokines provide potential targets for blockade and therapy. B lymphocytes can act as antigen-presenting cells in the synovium and also produce antibodies and secrete proinflammatory cytokines. B-cell depletion therapy has been shown to be beneficial in patients with RA [12]. The pannus begins to invade the joint cartilage and bone. Proliferation of the pannus leads to profound cartilage destruction, subchondral bone erosions, and periarticular ligament laxity. Cytokine-stimulated osteoclast activity also contributes to the erosions and periarticular osteoporosis seen in RA [13].

The search for specific immunologic mechanisms steered the discovery of autoantibodies in patients with RA. The role of these autoantibodies in RA pathogenesis is unclear. The long-recognized rheumatoid factor (RF), a polyclonal IgM directed against the Fc portion of IgG, is found in about two-thirds of RA patients. It binds to IgG-containing immune complexes and enhances immune responses. RF can be present in chronic infections such as hepatitis B and C and chronic inflammatory conditions. The sensitivity and specificity of RF for RA diagnosis are roughly 66% and 82%, respectively. The key autoantibodies associated with RA are those against citrullinated peptides (ACPA), as described before. The sensitivity of anti-CCP antibodies for the diagnosis of RA is 70%, but the specificity is 95%; thus, these antibodies are superior to RF in making a diagnosis of RA. Up to 35% of patients with a negative RF at presentation will test positive for anti-CCP antibodies. A small percentage of patients (10–15%) will remain seronegative (RF and CCP antibody–negative) throughout the course of the disease.

29.1.4 Diagnosis

The presentation and course of RA is variable. Typically, patients present with an insidious onset of symmetric joint pain, swelling, and morning stiffness worsening over several weeks. The severity and duration of morning stiffness often correlate with overall disease activity. Less common presentations include acute, rapidly progressive polyarthritis and, more rarely, monoarthritis. RA commonly involves the small joints of the hands (metacarpophalangeal [MCP] and proximal interphalangeal [PIP] joints), wrists, and feet (metatarsophalangeal [MTP] joints). Large joints can also be affected and include the shoulder, elbow, hip, knee, and ankle. Physical findings in RA involve the identification of symmetric joint inflammation early in the course of disease, as well as manifestations of joint destruction with chronic disease. Active synovitis is characterized by warmth, swelling, pain, and palpable effusions. Synovial proliferation is determined on physical examination by the presence of soft or rubbery tissue around the joint margins.

Severe RA can manifest with sequelae of systemic inflammation, especially in RF-positive patients. These include rheumatoid nodules and vasculitic skin ulcerations and ocular, pulmonary, cardiac, neurologic, and hematological abnormalities. Renal and gastrointestinal manifestations are rare. With early recognition and treatment of RA, the incidence of extra-articular manifestations diminishes. The American College of Rheumatology (ACR) classification criteria for the diagnosis of RA [14] were designed for inclusion of patients in clinical studies and not for clinical diagnosis. It is also pertinent to point out here that these criteria were formulated using data from hospitalized patients and thus are not very useful in the outpatient setting. Despite these caveats, these criteria are helpful to a physician unfamiliar with RA as a tool to evaluate a patient in the office. Most of the articular destruction occurs in the early years of disease, so it is important to diagnose and treat RA early [15,16].

29.1.5 Treatment

The goals of RA treatment are to alleviate pain, control inflammation, preserve and improve activities of daily living, and prevent progressive joint destruction. Medical treatment includes the use of nonsteroidal anti-inflammatory drugs (NSAIDs), disease-modifying antirheumatic drugs (DMARDs), and corticosteroids. Equally important in the management of RA is nonmedical treatment, including patient education, physical therapy, occupational therapy, orthotics, and, rarely, surgery.

DMARD therapy within three months of diagnosis is currently recommended because it can slow or arrest progression of the disease. DMARDs suppress immune-mediated inflammation by decreasing the activity of target cells (e.g., lymphocytes) or specifically targeting cytokine pathways [1]. Commonly used DMARDs include methotrexate, leflunomide, sulfasalazine, and hydroxychloroquine. When maximal dosing of one DMARD provides suboptimal disease control, the addition of other DMARDs often improves effectiveness [17,18]. Monitoring for RA disease activity and drug toxicity is important as well.

Biologic DMARDs targeting the inflammatory cytokine TNF-α include the fusion protein etanercept; the monoclonal antibodies infliximab, adalimumab, and golimumab; and the pegylated TNF receptor fusion protein certolizumab pegol. Other currently available biologic DMARDs include abatacept (a cytotoxic T-lymphocyte antigen 4 fusion protein that blocks T-cell costimulation by antigen-presenting cells); rituximab (a monoclonal antibody to CD20 that depletes B cells); the IL-1 receptor antagonist anakinra; and tocilizumab, a monoclonal antibody to the IL-6 receptor. Recently, an oral biologic, tofacitinib, which is a janus kinase 3 (JAK 3) inhibitor was approved for RA treatment.

Corticosteroids in low doses (e.g., prednisone, 5–10 mg) are extremely effective in promptly reducing the symptoms of RA and are useful in helping patients recover their previous functional status, but care should be taken to use the lowest effective dose. Corticosteroids are appropriate in patients with significant limitations in their activities of daily living or in an acute disease flare while awaiting the efficacy of the generally slower-acting DMARDs [1]. NSAIDs, including the selective cyclo-oxygenase (COX)-2 inhibitors, are commonly added to treatment regimens for pain relief and to decrease inflammation. NSAIDs do not prevent progression of bone and cartilage damage and should be used with caution because of their potential for gastrointestinal (GI) and renal toxicity and effects on cardiovascular risk, especially since RA patients may have an increased incidence of cardiovascular events [19].

29.1.6 Pharmacogenetics and Stratified Medicines in RA

The rapid development of biologic DMARDs has widened the choices of treatment for RA and have made therapy for it more effective but at the same time more expensive. Thus, in recent years there has been an interest in developing pharmacogenetic methods to predict an individual patient's response to a given DMARD.

This chapter highlights some of the major publications in the field of pharmacogenetics in RA and touches on the implications of this field for future research and clinical care. The pharmacogenetics of five major nonbiologic DMARDs—methotrexate, leflunomide, azathioprine, sulfasalazine, and hydroxychloroquine—and three biologic DMARDs—the tumor necrosis factor antagonists, rituximab, and tocilizumab—are reviewed in the following sections.

29.2 PHARMACOGENETICS OF NONBIOLOGIC DMARDS

29.2.1 Methotrexate

For over the past two decades, methotrexate (MTX) has been the first-line DMARD in RA because of its well-recognized efficacy and safety [20–22] However, the response among patients to it can be quite variable, ranging from 46 to 65% [23,24]. MTX's exact mechanism of action in RA remains unclear; however, current evidence suggests that it affects the intracellular folate and adenosine pathways.

MTX is actively transported into the cell by solute carrier 19A1 (SLC 19A1), also called the reduced folate carrier 1 (RFC1) (Figure 29.1). It is pumped out of the cell by members of the ATP-binding cassette (ABC) family of transporters, also known as multidrug-resistant transporters (MDRs), and multidrug resistance–associated proteins (MRPs) [25]. After entry into the cell, glutamate residues are added to MTX by folylpolyglutamyl synthase (FPGS); thus, MTX is polyglutamated ($MTXPG_n$). The enzyme gamma glutamyl hydrolase (GGH) can reverse this process. Polyglutamation keeps MTX within the cell, preventing drug efflux by the ABC transporters.

The ratio of GGH to FPGS has been suggested as a predictor of MTX treatment outcomes in human primary leukemia [26]. MTX-resistant cells show decreased FPGS activity [27–29]. MTXPGs inhibit dihydrofolate reductase (DHFR), which reduces dihydrofolate to tetrahydrofolate (THF) [30]. THF is converted to 5,10 methylene tetrahydrofolate (5,10-CH_2-THF) and subsequently to 5-methyl THF (5-CH_3-THF) by methylene tetrahydrofolate reductase (MTHFR). 5-methyl THF is a biologically active folate cofactor that functions as a 1-carbon donor for many important cellular reactions, including the conversion of homocysteine to methionine [31]. MTXPGs also inhibit thymidylate synthase (TYMS), which converts deoxyuridylate to deoxythymidylate in the de novo pyrimidine synthetic pathway [32].

MTX has numerous effects on the purine synthetic pathway. MTXPGs inhibit the enzyme aminoimidazole carboxamide ribonucleotide (AICAR) transformylase, which in turn results in intracellular accumulation of AICAR. AICAR and its metabolites can then inhibit two enzymes in the adenosine pathway—adenosine deaminase (ADA) and adenosine monophosphate (AMP) deaminase—which leads to intracellular accumulation of adenosine and adenine nucleotides. Subsequent dephosphorylation of these nucleotides results in increased extracellular concentrations of adenosine, which is a powerful anti-inflammatory agent [33].

Polymorphisms in genes encoding MTX transporters and enzymes in the folate and adenosine pathways inhibited by MTX have been studied in RA patients (Table 29.1).

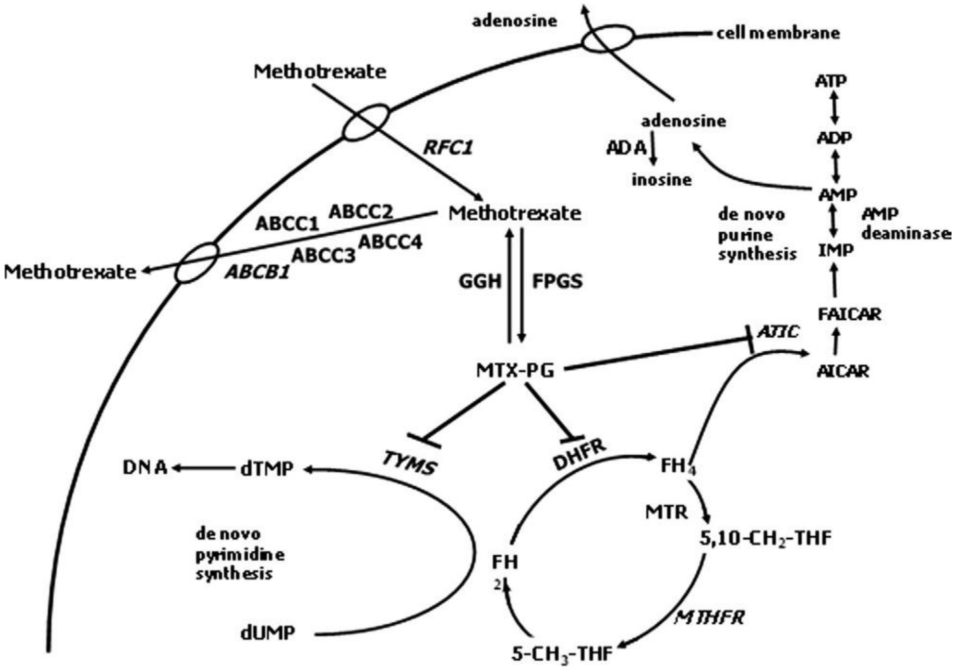

adenosine
cell membrane
Methotrexate
RFC1
adenosine
ADA
inosine
ATP
ADP
AMP
AMP deaminase
IMP
FAICAR
ATIC
AICAR
de novo purine synthesis
ABCC1
ABCC2
ABCC3
ABCC4
ABCB1
Methotrexate
Methotrexate
GGH
FPGS
MTX-PG
TYMS
DHFR
FH4
MTR
5,10-CH2-THF
MTHFR
5-CH3-THF
FH2
DNA
dTMP
dUMP
de novo pyrimidine synthesis

FIGURE 29.1 **Cellular pathway of methotrexate.** RFC1—reduced folate carrier 1; ABCB1, ABCC1-4—ABC transporters; GGH—γ-glutamyl hydrolase; FPGS—folylpolyglutamyl synthase; MTX-PG—methotrexate polyglutamate; TYMS—thymidylate synthase; d UMP—deoxyuridine monophosphate; d TMP—deoxythymidine monophosphate; DHFR—dihydrofolate reductase; FH$_2$_dihydrofolate; 5-CH$_3$-THF—5-methyl tetrahydrofolate; MTHFR—methylene tetrahydrofolate reductase; 5, 10-CH$_2$-THF—5, 10-methylene tetrahydrofolate; MTR—methyl tetrahydrofolate reductase; AICAR—aminoimidazole carboxamide ribonucleotide; FAICAR—10-formyl AICAR; ATIC—AICAR transformylase; IMP—inosine monophosphate; AMP—adenosine monophosphate; ADP—adenosine diphosphate; ATP—adenosine triphosphate; ADA—adenosine deaminase. Italicized genes have been targets of pharmacogenetic analyses in studies published so far. *Source: Modified from Davila and Ranganathan et al.* [34].

TABLE 29.1 Pharmacogenetics of Methotrexate

Gene	Role in MTX Pathway	Polymorphism	Postulated Effect of Polymorphism	Clinical Effects	Reference
RFC1	Active transport of MTX into the cell	80G>A	Increased transcriptional activity of *RFC1* gene; increased MTX entry into the cell	Increased response to MTX	[35,36]
ABCB1	MTX efflux from the cell	3435C>T	Increased MTX entry into the cell	Increased response to MTX	[37]
MTHFR	Regeneration of reduced folate; indirectly inhibited by MTX	677C>T	Thermolabile MTHFR with decreased activity; increased plasma homocysteine	Increased GI side effects	[38]
				Increased hepatic toxicity, GI toxicity, alopecia, stomatitis, and rash	[38–42]
				No effect on toxicity	[43]
				No effect on efficacy or toxicity	[44]

(*Continued*)

TABLE 29.1 Continued

Gene	Role in MTX Pathway	Polymorphism	Postulated Effect of Polymorphism	Clinical Effects	Reference
MTHFR		1298A>C	Decreased MTHFR activity; increased plasma homocysteine	Increased MTX efficacy	[41]
				Increased susceptibility to RA; decreased MTX toxicity	[43]
				No effect on efficacy or toxicity	[42,44,45]
DHFR	Reduction of dihydrofolate to tetrahydrofolate	473G>A 35289G>A	Decreased DHFR activity	No effect on MTX efficacy or toxicity	[45]
ATIC	Conversion of AICAR to 10-formyl AICAR; directly inhibited by MTX	347C>G	Decreased ATIC activity; increased AICAR accumulation; increased adenosine	Increased MTX efficacy	[36,41,46–48]
TYMS	Conversion of dUMP to dTMP; directly inhibited by MTX	5′UTR 28-bp repeat (TSER)	Increased TYMS enzyme activity	Decreased MTX efficacy	[49]
		3′UTR 6-bp deletion	Decreased TYMS mRNA stability and expression	Increased MTX efficacy	[49]
GGH	Reversal of polyglutamation of MTX	452C>T	Decreased GGH activity, decreased accumulation of intracellular long-chain MTX polyglutamates	Increased MTX efficacy	[50]
		3 G/3 G genotype	Increased GGH activity	Decreased response to MTX	[45]
				Increased bone marrow toxicity	[51]

Source: Modified from Davila and Ranganathan et al. [34].

29.2.1.1 Transporter Genes—RFC1/GGH/ABCB1

MTX is transported into the cell by RFC1. Polymorphisms in the *RFC1* gene, by altering gene expression, can modify MTX transport [52,53]. *RFC1* is a 27 kb gene located on chromosome 21 (21q22.3). An 80G>A polymorphism leading to substitution of arginine for histidine at codon 27 in the first transmembrane domain (TMD1) of the RFC1 protein and a 61- bp (base pair) repeat polymorphism in the *RFC1* promoter causing increased transcriptional activity of the gene have been described [54]. In a cross-sectional study conducted on 105 RA patients, the effect of the G80A single-nucleotide polymorphism (SNP) on response to MTX was examined. Patients in the top 25th percentile of MTX responders were identified using a visual analog scale (VAS) that measured the patients' and physicians' response to MTX. Patients homozygous for the *RFC* 80A/A genotype had a greater response to MTX compared to patients carrying the wild-type 80G/G. Patients homozygous for the A allele were three times more likely to be within the top 25th percentile of MTX responders (confidence intervals (CI) 1.3–8.4; $p<0.01$), compared to patients with the wild-type G allele. Thus, the *RFC* 80AA SNP may be a marker of amplified response to MTX in RA [55].

SNPs in the *GGH* promoter can impact both GGH expression [56] and MTX polyglutamation [50,57]. A 452C>T SNP leading to decreased GGH activity and accumulation of intracellular long-chain MTX polyglutamates [50] and a 401C>T promoter polymorphism also modifying intracellular MTXPG levels have been described [57].

There is also a *GGH*-16T>C polymorphism whose functional effects are unidentified. A Japanese study of patients with RA demonstrated that the presence of the *RFC* 80AA and *GGH*-401TT genotypes independently predicted MTXPG levels. Patients carrying the *RFC* 80AA genotype were 3.4-fold more likely to have MTXPG levels above the group median compared to patients with the 80GG or 80GA genotype (odds ratio (OR) 3.4, 95% CI 1.4–8.4; p = 0.007).

In contrast, patients with the *GGH*-401TT genotype were 4.8-fold (OR 4.8, 95% CI 1.8–13.0; p = 0.002) more likely to have MTXPG below the study median compared to those who carried the *GGH*-401CC or CT genotype [41]. Thus, both the *GGH*-401C>T and the *RFC* 80G>A SNPs influenced intracellular MTXPG levels and so may predict MTX response in RA [35,36,50,56].

29.2.1.2 *ABCB1*

ABCB1 is a 209 kb gene located on chromosome 7 (7q21.1). It codes for P-glycoprotein (P-gp), a membrane transporter important in the transport of a number of drugs, including MTX. SNPs in the *ABCB1* gene have been studied for their effects on P-gp expression [58]. The 3435C>T SNP is a synonymous SNP in exon 26 of the *ABCB1* gene. The 2677G>T SNP in exon 21, which results in substitution of alanine in position 893 by serine or threonine, is frequently linked to the 3435C>T SNP [59,60]. The impact of variations in *ABCB1* and/or P-gp expression on MTX efflux from the cell is uncertain. One study of human leukemic CCRF-CEM sublines suggested that higher P-gp expression may mediate MTX resistance [61]; however other studies did not support this [62,63].

In one study, genotyping was performed in 92 RA patients and 97 healthy controls for the 3435C>T polymorphism. Patients who had active RA (n = 62, nonresponders) after six months of treatment with MTX (7.5–15 mg/week) and prednisone (5–10 mg/day) and those who responded after six months of the same treatment (n = 30) were studied. The distribution of the *ABCB1* genotypes was similar between RA patients and controls. Patients with the 3435CC and 3435CT genotypes had active disease compared to patients with the 3435TT genotype (OR 2.89; CI 0.87–9.7; p < 0.05). Thus, the presence of the 3435T allele was associated with less active RA and a better response to MTX and prednisone [64].

29.2.1.3 *MTHFR/TYMS/DHFR*

MTHFR is a 19 kb gene located on chromosome 1 (1p36.3). Of several *MTHFR* polymorphisms that have been recognized, [65] two, 677C>T and 1298A>C, have been well studied for their influence on MTX's clinical properties. The 677C>T polymorphism leads to an alanine-to-valine substitution in the codon at nucleotide 677 of the *MTHFR* gene [66]. This change leads to the encoding of a thermolabile variant of MTHFR with resultant decreased enzyme activity. The 1298A>C polymorphism results in a glutamine-to-alanine substitution in the codon at nucleotide 1298. It leads to reduced MTHFR enzyme activity [67]. MTHFR is important in the generation of 5-methyl THF (see Figure 29.1). Because 5-methyl THF acts as the carbon donor for the remethylation of homocysteine to methionine, these two SNPs, by reducing MTHFR activity, can increase plasma homocysteine levels [68].

A number of studies have evaluated the effects of these SNPs on MTX response. One examined 105 patients with RA, 35 of whom were treated with MTX (7.5–15 mg/week), 34 with sulfasalazine (SSZ) (2–3 g/day), and 36 with MTX and SSZ. The 677C>T SNP was genotyped in all patients and their plasma homocysteine levels were measured. Patients on MTX had higher plasma homocysteine levels than those on SSZ alone, but those on both MTX and SSZ had the highest levels. Patients heterozygous for the 677C>T SNP had higher plasma homocysteine levels after one year than patients without the SNP. Patients homozygous for the SNP had a higher plasma homocysteine level at baseline, which did not change significantly. Elevated plasma homocysteine levels (17%, p < 0.05) were found in patients with GI side effects from MTX, such as nausea, abdominal pain, and discomfort compared to patients without side effects.

Combination therapy with MTX and SSZ led to the highest homocysteine levels and the most GI side effects. However, the presence of the C677T SNP was not directly associated with the occurrence of GI events. This suggests that MTX's toxic effects may be caused by elevated plasma homocysteine levels that may be exacerbated by the presence of the *MTHFR* 677C>T SNP [38].

In a Dutch study, 236 patients with RA on MTX were genotyped for the 677C>T SNP. MTX was initiated at 7.5 mg/week and titrated to a maximum dose of 25 mg/week. Patients were assessed for MTX toxicity and disease activity periodically. Of 236 patients, 122 (52%) did not have the SNP; 19 (8%) were homozygous; and 95 (40%) were heterozygous for the polymorphism. Patients who were homozygous and heterozygous for the 677C>T SNP had an increased risk of discontinuing MTX because of adverse events (relative risk (RR) 2.01; CI 1.09–3.70), particularly hepatotoxicity (RR 2.38; CI 1.06–5.34). Supplementation with folic or folinic acid reduced the risk of toxicity-related discontinuation of MTX both in patients with and without the SNP. The 677C>T SNP was not associated with MTX efficacy [39].

Berkun et al genotyped 93 RA patients treated with MTX (average dose 11.9 mg/week) and 377 healthy controls for the 677C>T and 1298A>C polymorphisms in a cross sectional study and assessed patients for RA disease activity and MTX toxicity. Serum folate and plasma homocysteine levels were measured. The prevalence of the 1298CC genotype (24.7%) was higher in RA patients than controls

(12.8%) and this was significant ($p < 0.001$). There were no associations between the 1298A>C SNP and MTX efficacy. Homozygotes for the 1298C allele appeared to be protected from MTX toxicity; 33% did not experience toxicity, only 9.1% had adverse reactions ($p = 0.035$). In contrast, patients with the 1298AA genotype were five times more likely to develop toxicities than those with the CC genotype (OR 5.24; CI 1.38–20). Also, patients carrying the C allele had higher plasma homocysteine levels than patients with the AA or AC genotype, and this was not influenced by serum folate levels. Because the 677C>T polymorphism had no effect on MTX toxicity or efficacy in this study, homozygosity for the 1298C allele may increase susceptibility to RA, but also may protect from MTX toxicity presumably via a homocysteine-dependent mechanism [43].

In another study, 106 RA patients who had been treated with MTX were assessed retrospectively for MTX efficacy and toxicity. The dose ranged from 2.5 mg/week to 12.5 mg/week. All patients were genotyped for the 677C>T and 1298A>C SNPs. Direct assessment of MTX efficacy (as measured by disease activity scores) was not possible because of the retrospective design. However, patients carrying the 1298A>C polymorphism (homozygous or heterozygous) were more likely to be on lower doses of MTX compared to those without the polymorphism (RR 2.18; CI 1.17–4.06; $p < 0.05$). Patients carrying this polymorphism also showed improvement in their inflammatory markers such as erythrocyte sedimentation rate (ESR) and C reactive protein (CRP), suggesting an increased clinical efficacy of MTX ($p < 0.05$). However, the presence of the 677C>T polymorphism was associated with increased MTX toxicity (RR 1.25; CI 1.05–1.49; $p < 0.05$) but not with MTX efficacy. Thus, the 677C>T polymorphism appeared to be a marker for MTX toxicity and the *1298A>C* polymorphism for MTX efficacy [40].

More recent studies have yielded varied results. Three confirmed an effect of the 677C>T polymorphism on MTX efficacy. However, results from these studies were conflicting, with the T allele being a marker of both decreased and increased MTX efficacy in US [46] and Polish [69] cohorts, respectively, and the C allele being a marker of increased MTX efficacy in a Dutch cohort [41]. Eight recent studies showed an effect of the T allele on MTX toxicity. Four examined Asian patients, including Japanese, [40,70] Korean, [71] and Chinese; [72]; others included Dutch [73], US, [47,74], and Spanish [75] cohorts. Two meta-analyses have also yielded contrasting results. One found an association between the 677C>T polymorphism and MTX toxicity, but no such association for the 1298A>C variant [42]. However, another (which included 1514 patients with RA) found no association between either of these polymorphisms and MTX toxicity and efficacy [44].

TYMS is an important enzyme in the de novo synthesis of pyrimidines. It converts deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP) and is a direct target of polyglutamated MTX. The *TYMS* gene is a 15 kb gene located on chromosome 18 (18p11.32). A polymorphic tandem 28-bp repeat sequence has been described in the 5′ untranslated region (TSER) of the *TYMS* gene, with a variable number of repeat elements [76]. *In vitro* studies have shown that TYMS mRNA expression and enzyme activity increases with an enhancing number of these repeat sequences [76–78]. Patients homozygous for the triple repeat allele (*TSER*3/*3* or *3R*) have higher TYMS mRNA expression compared to patients homozygous for the double repeat allele (*TSER*2/*2* or *2R*) [78,79]. An additional G to C substitution within the 3R allele further diversifies 3R into 3RC and 3RG; the C allele abolishes a critical residue in the upstream stimulatory factor (USF-) –binding site, lowering TYMS activity in 3RC carriers [80]. Deletion of a 6-bp sequence at nucleotide 1494 in the 3′ untranslated region (3′UTR) of TYMS has also been described and may be associated with decreased TYMS mRNA stability and expression [45,81].

In a retrospective study, 167 patients with RA, of whom 115 were treated with MTX, were genotyped for the following polymorphisms—*TYMS* 5′UTR enhancer repeat (*TSER*), 3′UTR deletion, *MTHFR* 677C>T, and 1298A>C. The mean weekly MTX dose in this study was 5.7+/-2.3 mg per week. Both MTX-treated and -untreated groups displayed similar frequencies of these SNPs. The *TYMS* and *MTHFR* polymorphisms were not associated with toxicity, although a significant percentage of patients (45%) on MTX experienced adverse events. Weekly MTX dose (rather than standardized disease activity measures) was used as a marker of efficacy in this study. A dose of >6 mg/week was considered indicative of less efficacy and <6 mg/week was considered indicative of greater efficacy. Homozygotes for the *TSER*2* allele (*TSER*2/*2*) required lower doses of MTX (i.e., had greater efficacy) than homozygotes for the *TSER*3* allele ($p = 0.033$). The *TYMS* 3′UTR deletion and *MTHFR* polymorphisms had no effects on MTX efficacy. The researchers speculated that the repeat enhancer *TSER*3* polymorphism may lead to decreased MTX efficacy by increasing TYMS mRNA expression. This study also put forward the idea that *MTHFR* polymorphisms may not influence MTX toxicity or efficacy [49].

DHFR reduces dihydrofolate to tetrahydrofolate in the intracellular folate pathway. It is directly inhibited by polyglutamated MTX and encoded by the *DHFR* gene, a 28 kb gene located on chromosome 5 (5q11.2–q13.2). Two hundred and five MTX-treated RA patients were studied for *DHFR* gene polymorphisms. MTX was started at 7.5 mg/week and increased to 15 mg/week (with folic acid supplementation) after four weeks based on response to the drug. MTX efficacy and toxicity were assessed from time to time. Genotyping for the *MTHFR* 677C>T, *MTHFR* 1298A>C, *DHFR* -473G>A, *DHFR*

35289G>A, and *RFC* 80 G>A SNPs was performed. At six months, patients carrying the *MTHFR* 1298AA and *MTHFR* 677CC (wild-type) genotypes showed a greater response to MTX compared to patients carrying the heterozygous or homogenous genotype (OR 2.3; CI 1.18–4.41 and OR 2.73; CI 1.03–7.26, respectively). Haplotype analysis for the *MTHFR* 1298A and 677C SNPs revealed that patients with two copies of the haplotype had greater improvement than those with one or no copies (OR 3.0; CI 1.4–6.4). Patients homozygous and heterozygous for the *MTHFR* 1298 A>C SNP (*MTHFR* 1298AC+CC) had an increased number of overall adverse drug events at three and six months (OR 2.55; CI 1.20–5.41 and OR 2.5; CI 1.32–4.72, respectively). Because the *RFC* and *DHFR* SNPs were not associated with MTX toxicity or efficacy, patients with the wild-type *MTHFR* alleles (*MTHFR* 1298AA and 677CC) responded better to MTX, while those with the 1298C allele had an increased risk for MTX toxicity [41].

Thus, based on the literature just cited (Table 29.1), the 677C>T SNP in *MTHFR* appears to have effects on MTX toxicity, presumably through its effects on homocysteine metabolism, and on MTX efficacy [41]. The effects of the 1298A>C polymorphism on MTX are inconsistent, with data suggesting that it may increase [40] or decrease [41] patients' response to MTX and possibly protect them from MTX toxicity [43]. This may stem from the fact that these SNPs perhaps have effects other than that on homocysteine metabolism that may influence response to MTX. The retrospective design of most of these studies may not have allowed an accurate assessment of MTX effects, especially adverse effects. Although one of the studies [49] concluded that *MTHFR* SNPs did not affect MTX efficacy or toxicity, it is worth pointing out that the doses of MTX used in this study were small (6 mg/week as opposed to standard doses of 15–20 mg/week), which may have influenced the study results. Also, standardized measures of disease activity were not used in this study, which may have further impacted the results.

29.2.1.4 ATIC /CCND1

AICAR transformylase (ATIC) converts AICAR to 10-formyl AICAR and is directly inhibited by MTX (see Figure 29.1). This leads to accumulation of AICAR and adenosine, a purine with anti-inflammatory properties. Adenosine may be an important mediator of the anti-inflammatory effects of MTX [82]. *ATIC* is a 37 kb gene located on chromosome 2 (2q35). The *ATIC* 347C>G SNP leads to a threonine-to-serine substitution in codon 2, which may cause a decrease in ATIC's enzymatic activity and affect AICAR accumulation and adenosine release. In 108 RA patients on MTX at a dose of 14 mg/week (range 5–25 mg/week), the combined effects of the *ATIC* 347C>G SNP, *TSER*2*, and the *RFC* 80G>A polymorphism on MTX efficacy was assessed. Red blood cell long-chain MTX polyglutamate (MTXPG) concentrations were measured and a pharmacogenetic index was calculated from the sum of homozygous variant genotypes (*RFC1* 80AA, *ATIC* 347GG, *TSER*2/*2*). The allelic frequency of the *ATIC* 347G variant was 37%. Patients were categorized as MTX responders or MTX nonresponders using a visual analog scale (VAS).

Eighteen patients who were carriers of the *ATIC* 347GG genotype had fewer swollen joints (1.9 ± 0.6 versus 4.5 ± 0.6 [p = 0.06]) and a lower score for physician's assessment of patients' response to MTX (1.8 ± 0.3 versus 2.8 ± 0.2 [p = 0.02]) compared to 90 patients who were carriers of the C allele (*ATIC* 347CC, n = 47; *ATIC* 347CG, n = 43) [36]. Among other studies, only one demonstrated the *ATIC* 347C allele to be associated with MTX efficacy in a Dutch cohort [41]. Four studies showed an association of the *ATIC* 347G allele with MTX toxicity in US [46,47], Dutch [41], and Slovenian cohorts [48].

G1/S-specific cyclin-D1 is a protein that in humans is encoded by the *CCND1* gene, whose activity is required for cell-cycle G1/S transition [83]. The 870A>G substitution resides at the mRNA splicing site [84]. The A allele preferentially encodes the transcript lacking exon 5, leading to increased levels and longer half-life of cyclin D1 [85]. Elevated levels or expression of cyclin D1 protein (coded by the A allele) have been found in a variety of cancers, including breast cancer, head and neck cancer, non-small-cell lung cancer, and mantle cell lymphomas [86]. In a human fibro sarcoma HT1080 cell line, CCND1 over expression modified sensitivity to MTX [86].

Jekic et al. genotyped 184 RA patients treated with MTX for selected polymorphisms in the *GGH* (−354G>T and 452C>T), *CCND1* (870A>G) and *TYMS* (variable number of tandem repeats, or VNTR, and G-to-C substitution of the triple repeat, 3R allele) genes. Based on the European League Against Rheumatism (EULAR) response criteria—a standardized measure of disease activity in RA—146 RA patients (79.3%) were classified as responders and 38 (20.7%) as nonresponders after six months of MTX therapy. There was no difference between MTX responders and nonresponders in the frequency of polymorphisms in the *GGH* and *CCND1* genes or *TYMS* VNTRs. However, when *TYMS* was analyzed with respect to VNTRs and the 3R G-to-C substitution, a higher frequency of the 3G/3G genotype (G allele) was found in MTX nonresponders when compared to other genotypes (p = 0.02, OR 5.4, 95% CI 1.0–21.1) [51]. Eight of 184 patients developed bone marrow toxicity, and all of those eight carried the *GGH*-354GG genotype. The -354 T allele has been shown to correlate with increased *GGH* gene expression [56]. For this reason, the investigators proposed that the -354 G allele possibly results in reduced GGH levels and higher cellular concentrations of MTXPG, leading to increased toxicity [51].

29.2.2 Pharmacogenetics of Leflunomide

Leflunomide is an isoxazole derivative used as a DMARD in the treatment of RA. It is a prodrug that is metabolized in the plasma and intestinal mucosa to an active, open-ring metabolite, teriflunomide (also known as A77 1726), that causes noncompetitive and reversible inhibition of dihydroorotate dehydrogenase (DHODH) [87]. DHODH is critical to de novo pyrimidine synthesis, and the effect of leflunomide treatment is a decrease in lymphocyte proliferation and thereby decreased inflammation (Figure 29.2). Polymorphisms that may affect the response to leflunomide are discussed in this section, and summarized in Table 29.2.

29.2.2.1 DHODH

The human *DHODH* sequence frequently contains a missense polymorphism in the first exon (19A>C), which leads to a lysine-to-glutamine amino acid substitution in the N-terminal region of the polypeptide. This segment controls the insertion of the polypeptide into the mito chondrial inner membrane and thus influences the cellular effects of the enzyme [34].

Pawlik et al. [37] conducted a study on 147 patients (123 women, 24 men; mean age 52.8 ± 11.03 years and of Caucasian origin) diagnosed with RA and treated with leflunomide 20 mg/day. Genotyping was performed to analyze the 19C>A SNP in exon 1 of the *DHODH* gene. Remission was more frequent in patients who carried the C rather than the A allele, with no significant association between either of the alleles and adverse events. Conversely, in a retrospective study of 105 patients with RA, of whom 62 tolerated leflunomide treatment well, and 43 patients who discontinued treatment owing to leflunomide toxicity, this polymorphism was associated with a 6.8-fold increased risk of overall leflunomide-related adverse events, including GI, mucosal, and hepatic toxicity [88].

29.2.2.2 ABCG2

ATP-binding cassette subfamily G, member 2 (ABCG2) is a protein encoded by the *ABCG2* gene [89]. ABCG2 is involved in the disposition of various chemotherapeutic agents as an efflux transporter. It is highly expressed in the gastrointestinal tract and liver, and governs the absorption, distribution, and excretion of a wide variety of clinically important drugs, including leflunomide [90]. Two common *ABCG2* genetic variants are 34G>A and 421C>A.

Kim et al. [91] analyzed the pharmacokinetic characteristics of A771726, according to *ABCG2* genotypes, in 24 healthy study participants following a single leflunomide

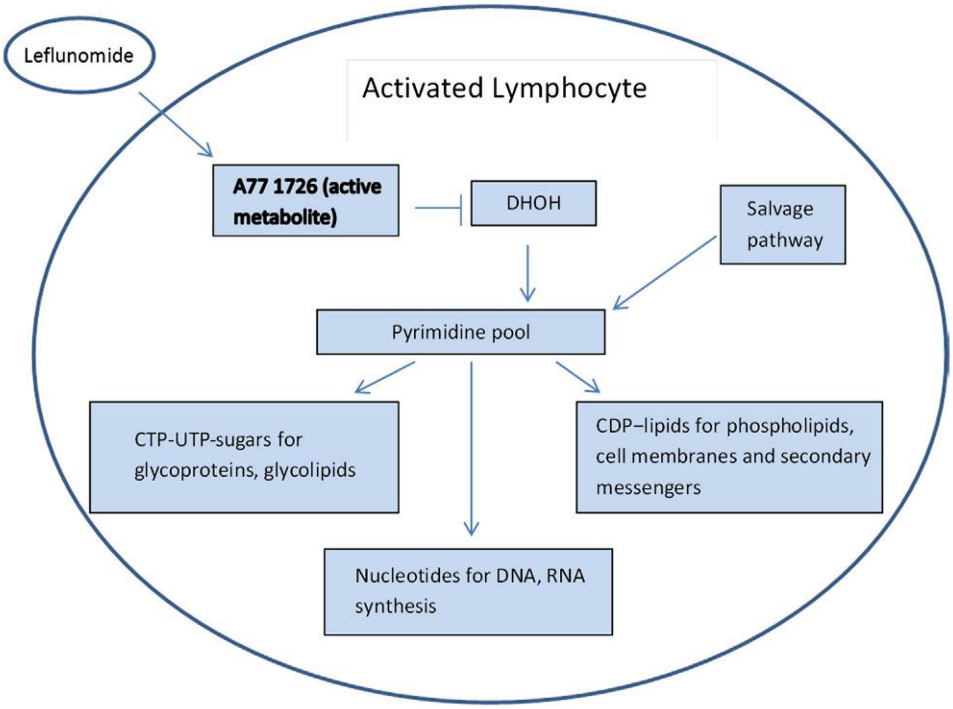

Leflunomide
Activated Lymphocyte
A77 1726 (active metabolite)
DHOH
Salvage pathway
Pyrimidine pool
CTP-UTP-sugars for glycoproteins, glycolipids
CDP–lipids for phospholipids, cell membranes and secondary messengers
Nucleotides for DNA, RNA synthesis

FIGURE 29.2 Metabolism of leflunomide. The active metabolite of leflunomide (A77 1726), whose CYP enzyme–mediated production occurs in plasma and intestinal mucosa, is taken up by activated lymphocytes, wherein it inhibits the action of DHODH. De novo pyrimidine synthesis is thus inhibited, leading to decreased lymphocyte proliferation. Polymorphisms in DHODH and CYP enzyme genes are thought to affect leflunomide toxicity. CYP—cytochrome P450; DHODH—dihydroorotate dehydrogenase. *Source: Modified from Davila and Ranganathan et al.* [34].

TABLE 29.2 Pharmacogenetics of Leflunomide

Gene, Product, and Role	Polymorphism	Cellular Effects of Polymorphism	Clinical Effects (Reference)
DHODH, DHODH, key enzyme in de novo pyrimidine synthesis	19 C>A	Alters N-terminal region of DHODH, affecting insertion into mitochondrial membrane	Associated with increased efficacy [37]
			Associated with increased toxicity [88]
			Not associated with toxicity [97]
ABCG2, ABCG2, governs absorption, distribution, and excretion of drug	421C>A	Influences bioavailability and oral clearance	Associated with increased C_{max} and decreased oral clearance [91]
			Associated with increased toxicity [97]
ESR1, estrogen receptor, interferes with down-modulation of cytokines	rs9340799 AA rs2234693 TT	Alters estrogen receptor expression	Associated with increased efficacy [93]
CYP1A2, CYP enzyme, involved in activation of leflunomide	CYP1A2*1F CC	Increases formation of active metabolite, leading to higher drug levels	Associated with increased toxicity [95]
	CYP2C19*2		Increased oral clearance [96]
	CYP2C19		Associated with increased likelihood of discontinuing leflunomide [97]

Source: Modified from Davila and Ranganathan et al. [34].

dose of 20 mg. *ABCG2* 421C>A substantially influenced the pharmacokinetics of A771726. A771726 C_{max} was 30% higher, the area under the concentration-time curve (AUC) was 83% larger, and oral clearance was 41% lower in 421C>A carriers than in noncarriers.

29.2.2.3 Estrogen Receptor Polymorphisms

Because previous studies [92] indicated that women have a poorer response to leflunomide treatment than men, Dziedziejko et al. [93] looked at the effect of estrogen receptor polymorphisms on leflunomide therapy. Two estrogen receptors—α and β—are known, and the genes *ESR1* and *ESR2* that encode them are responsible for transducing extracellular signals into transcriptional responses. In this prospective study, common SNPs in *ESR1* (rs9340799A>G, rs2234693T>C) and *ESR2* (rs4986938G>A, rs1256049G>A) were examined in 115 women with RA, mean age 54.1 ± 11.0 years, treated with leflunomide (20 mg/day). The *ESR1* rs9340799 AA and rs2234693 TT genotypes were associated with a better response to leflunomide after 12 months of therapy—improvement of ESR, patient's global assessment of disease activity on VAS and DAS28 (disease activity score, a standardized measure of disease activity in RA) values. There was no such association with *ESR2* genotypes and haplotypes [93]. A possible explanation for this association is that these polymorphisms may influence estrogen receptor expression and thus alter response to leflunomide.

29.2.2.4 Cytochrome P450 System

The cytochrome P450 (CYP) system, particularly the enzyme encoded by *CYP1A2*, converts leflunomide to its active metabolite [94] A prospective study from central Europe of 105 patients with RA identified an association between the *CYP1A2*1F* polymorphism and leflunomide toxicity [95]. Another prospective study by the same group, of 67 patients with RA and 4 with polyarthritis resembling RA and psoriasis treated with leflunomide, investigated the association of SNPs in *CYP1A2* and *CYP2C19* with A77 1726 pharmacokinetics. The oral clearance of A77 1726 was found to be 71% higher in carriers of the *CYP2C19*2* 681G>A polymorphism compared with noncarriers. There was no impact of steady-state plasma concentrations of A77 1726 on the occurrence of adverse events from the drug [96].

In a recent retrospective study from Australia, 78 RA patients on leflunomide for at least 12 months were genotyped for the following SNPs: *CYP2C19*2* (rs4244285), *CYP2C19*17* (rs12248560), *ABCG2*421C>A (rs2231142), *CYP1A2*1F* (rs762551), and *DHODH*19C>A (rs3213422). Polymorphisms were chosen based on the potential importance of encoded proteins on steady-state teriflunomide blood levels via an effect on leflunomide metabolism (*CYP2C19*2* [*CYP2C19* 681G>A, rs4244285] and *CYP2C19*17* [*CYP2C19*, 608C>T, rs12248560]) or on biliary secretion of teriflunomide (*ABCG2*421C>A, rs2231142) [91].

Polymorphisms previously linked to leflunomide toxicity—*CYP1A2*1F* (*CYP1A2* 163C>A, rs762551) [95,96] and *DHODH* (*DHODH* 19C>A, rs3213422) [88]—were also studied. *CYP2C19*2* is the loss-of-function allele, while *CYP2C19*17* is the gain-of-function allele. Poor metabolizers of leflunomide carry two loss-of-function alleles, intermediate metabolizers carry one loss-of-function and one wild-type allele, extensive metabolizers carry two wild-type alleles, and ultrarapid metabolizers carry either one or two gain-of-function alleles and no loss-of-function alleles.

After adjusting for current leflunomide dose, triple therapy (with methotrexate, sulfasalazine, and hydroxychloroquine), and positive rheumatoid factor status at diagnosis, the *CYP2C19* genotype was significantly associated with the likelihood of discontinuing leflunomide, with poor and intermediate metabolizers discontinuing it more frequently (adjusted hazard ratio = 0.43 for each incremental change in genotype, 95%; CI 0.24–0.79; p = 0.006) [97]. Previously observed associations between *CYP1A2* [95] and *DHODH* genotypes and toxicity [88] were not apparent, but there was a trend for the *ABCG2* SNP to be associated with drug cessation due to diarrhea.

29.2.3 Pharmacogenetics of Azathioprine

Azathioprine (AZA) is used in the treatment of several rheumatic diseases, particularly systemic lupus erythematous (SLE) and RA. About 10–30% of RA patients discontinue AZA because of side effects [98]. AZA is a prodrug that after oral intake is converted into 6-mercaptopurine (6-MP), an active purine antimetabolite that affects the purine de novo synthetic and salvage pathways (Figure 29.3). 6-MP is converted by hypoxanthine phosphoribosyl transferase (HPRT) to cytotoxic thioguanine nucleotides (6-TGN) via the intermediary metabolite, 6-thio-inosine monophosphate (6-TIMP). Inosine monophosphate is phosphorylated to inosine triphosphate (ITP), a toxic metabolite, and this process can be reversed by inosine triphosphate pyrophosphatase (ITPase, encoded by *ITPA*). ITPase deficiency can result in the accumulation of toxic ITP. ITPase-deficient individuals treated with AZA can develop toxicity because of accumulation of thio-ITP [99]. By a parallel pathway, 6-MP can be deactivated by thiopurine methyltransferase (TPMT) to 6-methylmercaptopurine (6-MMP) or by xanthine oxidase (XO) to thiouric acid (TU). Thus, a relative or absolute deficiency of TPMT leads to accumulation of cytotoxic TGN and significantly increased AZA toxicity.

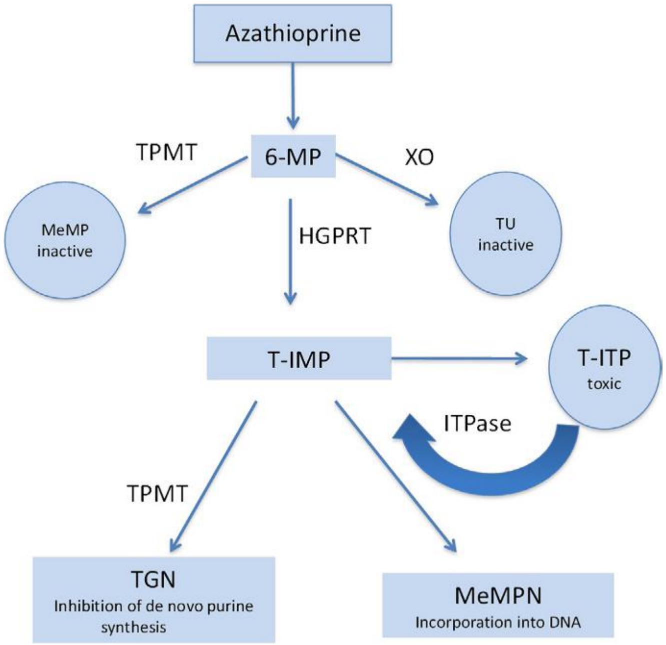


FIGURE 29.3 **Thiopurine drug metabolism.** TPMT—thiopurine methyltransferase; XO—xanthine oxidase; HPRT—hypoxanthine phosphoribosyl transferase; 6-TIMP—6-thiosine monophosphate; 6-MMP—6-methylmercaptopurine; 6-TGN—6-thioguanine nucleotides. *Source: Modified from Davila and Ranganathan et al.* [34].

The most common adverse effects of AZA are hematologic and gastrointestinal.

29.2.3.1 TPMT

The *TPMT* gene is a 26 kb gene located on chromosome 6 (6p22.3). TPMT activity in erythrocytes varies between high, intermediate, and low or none and is controlled by the gene's allelic variants. Population studies have shown that approximately 90% of the population has high activity, 10% has intermediate activity, and 0.3% has little to no activity [100].

One of three common allelic variants of the *TPMT* gene—*TPMT*2*, *TPMT*3A*, or *TPMT*3C*—is found in 80–95% of patients with low TPMT activity [101–103]. Ethnicity seems to influence the frequencies of these allelic variants [104,105]. Patients with low TPMT activity require lower AZA doses to avoid toxicities, as standard doses may lead to significant toxicities [106].

Sixty-eight patients with rheumatic disease on AZA (2–3 mg/kg/day) were genotyped for *TPMT*2* and *TPMT*3A* alleles. All patients were assessed for side effects from AZA, such as leucopenia, liver function abnormalities, and GI intolerance. Six (9%) were heterozygous for *TPMT*3A*; five of those discontinued AZA within four weeks of starting the medication because of hematologic toxicity [107]. In another study 40 RA patients on AZA were genotyped for the *TPMT* alleles. AZA dose varied from 0.7 mg/kg to 1.4 mg/kg/day. Six out of 40 patients discontinued AZA because of toxicity. Three of the six with severe GI toxicity were heterozygous for the *TPMT*3A* allele, while the remainder possessed the wild-type *TPMT* allele. The association between the *TPMT* allele and AZA toxicity was significant ($p=0.018$). The positive predictive value for toxicity in a *TPMT* polymorphism carrier was 60% in this particular study [108].

One case report described a SLE patient who presented with pancytopenia, sepsis, typhlitis, and disseminated intravascular coagulopathy after a short period of AZA therapy. The patient had low TPMT activity due to the *TPMT*3C* genotype [109]. A meta-analysis of 67 studies examined whether patients with intermediate TPMT activity were at increased risk of myelosuppression when taking thiopurine medications. It included all primary studies of patients on a thiopurine medication with either genotypic or phenotypic testing for TPMT activity, and reported cases of hematological adverse reactions. The search was not restricted to a specific disease or condition. Patients with two *TPMT* variant alleles who were TPMT-deficient had a substantial increase in their risk of myelotoxicity (86% of deficient patients developed myelosuppression). Patients heterozygous (i.e., with one variant allele) for any of the *TPMT* variant alleles that lead to intermediate TPMT activity were also at high risk for drug-induced myelosuppression compared to those with wild-type alleles (OR 4.19; 95% CI 3.20−5.48) [110].

29.2.3.2 ITPA

AZA is frequently used to treat inflammatory bowel disease. As mentioned earlier, ITPase deficiency can result in the accumulation of toxic ITP. In a prospective study of patients with Crohn's disease taking AZA, dropouts during the first two weeks because of adverse events were more frequent carriers of the *ITPA* 94C>A allele [111]. In a more recent study of patients with inflammatory bowel disease, 40 of 160 on AZA were found to have decreased ITPA activity (below the lower limit of the reference range <122 μmol/(gHb·h)). ITPA activity was measured in erythrocyte lysates by a high-performance liquid chromatography procedure based on conversion of ITP to IMP; it was expressed as micromoles of IMP formed per gram of Hb per hour (μmol/(gHb·h)). Patients with decreased ITPA activity were separated into quartiles, resulting in three additional thresholds (37.3, 59.9, and 89.2 μmol/(gHb·h)). Adverse events were reported for 88 of the 160 patients (55%). Patients with ITPA activity <89.2 μmol/(gHb·h) developed leukopenia more often than did patients with higher enzyme activity (OR 3.27; 95% CI 1.21−8.82), whereas individuals with very low ITPA activity (<37.3 μmol/(gHb·h)) had elevated liver enzymes more frequently compared with those with ITPA activity >37.3 μmol/(gHb·h) (OR 5.0; 95% CI 1.35−18.57). The only significant genotyping association was seen for the *ITPA* 94C>A polymorphism and leukopenia [112].

In contrast, an earlier study by Gearry et al. failed to demonstrate a significant association between the presence of *ITPA* alleles and AZA-related toxicity. In this study, genotyping was performed for the 94C>A missense mutation in *ITPA*, *TPMT*2*, and *TPMT*3* in 147 patients with inflammatory bowel disease. About 50% of patients had adverse events from AZA. There was no significant difference in the frequency of the *ITPA* alleles between patients who had an adverse event versus those who did not (16/146 versus 16/148, $p=0.56$). The *ITPA* 94C>A polymorphism was not found to be associated with adverse events such as rash or pancreatitis or flu-like symptoms [113].

29.2.3.3 GST

The conversion of AZA to the active metabolite 6-MP was considered to be a nonenzymatic reaction until recently. Studies in the last few years have shown that glutathione S transferases (GSTs) may be involved in this conversion [114,115]. One recent study demonstrated that patients with inflammatory bowel disease who carried the *GST-M1* null genotype required a lower AZA dose to generate 6-TGN compared to those with the *GST-M1* wild-type genotype. Other *GST* genotypes did not show a substantial effect on AZA metabolism [116].

29.2.3.4 Other Purine Pathway Enzymes

Some studies have looked at the association between the activity of enzymes other than TPMT in the purine pathway and AZA toxicity. In one study, TPMT, HPRT, 5′-nucleotidase, and purine nucleoside phosphorylase activity in the red blood cells (RBC) of 33 RA patients on AZA (≈2 mg/kg/day) and 66 controls was measured. Fourteen RA patients with low ($p = 0.004$) and seven patients with intermediate TPMT activity ((RR 3.1) developed AZA toxicity when compared to patients with normal TPMT activity [98]. Another study measured TPMT activity in three RA patients who had experienced AZA-induced hematologic toxicity and 16 RA patients without AZA toxicity. Two patients with AZA-induced hematologic toxicity were TPMT-deficient, one partial and the other complete. In the third patient, 5′-nucleotidase activity was found to be well below the lowest level observed in the control subject [117]. It should be pointed out that the sample sizes in these studies were small and genotyping was not done.

Both *TPMT* genotyping and measurement of TPMT activity in RBC have been studied in predicting and preventing AZA toxicity. Current evidence supports that TPMT deficiency, either partial or complete, can cause increased toxicities from AZA. Clearly, large prospective studies are needed to confirm the observations from the smaller studies just described (Table 29.3). There is some evidence for *ITPA* variants influencing AZA toxicity in inflammatory bowel disease, but there are no studies to date on this polymorphism in RA. Of note, *TPMT* genotyping is available to clinicians to screen patients prior to initiation of AZA, and it is the first commercially available assay for pharmacogenetic testing in rheumatology.

29.2.4 Sulfasalazine

Sulfasalazine (SSZ) is another DMARD often used in the treatment of RA. It is estimated that 20–30% of RA patients on SSZ report adverse drug reactions. The major adverse drug events of SSZ are gastrointestinal and hematologic.

29.2.4.1 NAT2

SSZ is a combination of sulfapyridine and 5-aminosalicylic acid (5-ASA). After ingestion, SSZ is split into these two compounds by the colonic bacteria. 5-ASA remains in the large bowel, while most of the sulfapyridine is completely absorbed and undergoes acetylation, hydroxylation, and glucoronidation in the liver. The enzyme N-acetyltransferase 2 (NAT2) acetylates sulfapyridine into N-acetylsulfapyridine (Figure 29.4). *NAT2* is a 9 kb gene located on chromosome 8 (8p22). Its polymorphisms may modify the acetylator phenotype of an individual (slow versus fast) and thus have effects on an individual's vulnerability to SSZ toxicity. SSZ toxicities such as abdominal discomfort, nausea, rash, and headaches are found more commonly in slow than in fast acetylators [118,119].

A few studies have evaluated the effects of *NAT2* polymorphisms on SSZ toxicity in RA patients. One retrospective study evaluated 144 RA patients on SSZ at a dose range of 500–1500 mg/day. *NAT2* genotyping was carried out in all patients. Slow acetylators lacking the wild-type *NAT2*4* allele had a higher frequency of adverse events (63%) compared to fast acetylators with at least one *NAT2*4* allele (8%). The association between the *NAT2* genotype and SSZ toxicity was significant (OR 7.73; CI 3.54–16.86;

TABLE 29.3 Pharmacogenetics of Azathioprine

Polymorphism	Amino Acid Change in Enzyme	Population Prevalence	Effect on Enzyme Activity	Biochemical Effect	Clinical Effects	Reference
*TPMT*2* 238G>C	Alanine to proline	0.2–0.5%	Low to intermediate due to enhanced degradation of enzyme	Decreased methylation of drug to inactive compounds	Hematologic and GI toxicity	[107]
*TPMT*3A* 460G>A, 71A>9G	Alanine to threonine and tyrosine to cystine, respectively	3.2–5.7%	As above	As above	Hematologic toxicity	[107,108]
*TPMT*3C* 719A>G	Tyrosine to cystine	0.2–0.8%	As above	As above	Hematologic toxicity	[109]

Source: Modified from Davila and Ranganathan et al. [34].

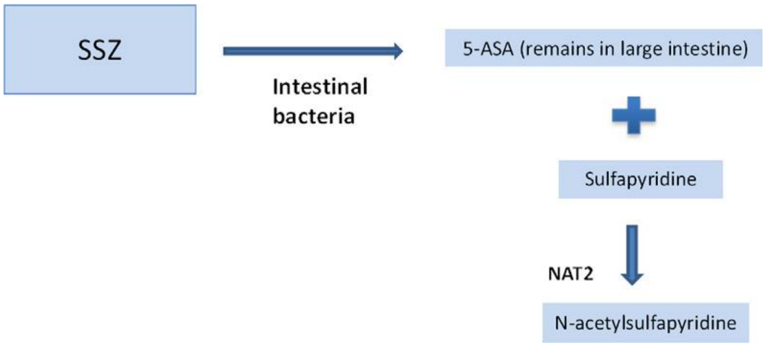


FIGURE 29.4 **Sulfasalazine metabolism.** SSZ—sulfasalazine; 5-ASA—5-aminosalicylic acid; NAT 2—N-acetyltransferase 2 NAT2 acetylates sulfapyridine into 5-acetylsulfapyridine. *Source*: *Modified from Davila and Ranganathan et al.* [34].

$p < 0.001$). In fact, 25% of the slow acetylators had to be hospitalized for their toxicities [120]. In another study, 114 patients with inflammatory bowel or joint disease treated with SSZ (mean dose 2 gm/day) were genotyped for 5 allelic variants—*NAT2*5A*, *NAT2*5B*, *NAT2*5C*, *NAT2*6*, and *NAT2*7* (encoding slow acetylator status) and the wild-type *NAT2*4* allele (encoding rapid acetylator status). Twenty-seven of 39 patients (69%) who developed agranulocytosis within three months of starting treatment with SSZ were slow acetylators compared to 34 of 75 patients (45%) who received SSZ without a hematological adverse event (OR 2.7; $p = 0.002$). *NAT2*6* alleles were found to be more frequent among patients who had agranulocytosis (36%) compared to patients without it (23%) ($p = 0.033$) [121].

In a more recent investigation, researchers performed two bioavailability studies under comparable conditions with 24 healthy subjects. Plasma levels of sulfapyridine and acetylsulfapyridine were determined after oral intake of enteric-coated formulations of sulfasalazine (500 mg and 1000 mg, respectively). The *NAT2* genotype was analyzed in parallel for all subjects. The terminal elimination half-life of sulfapyridine as well as of acetylsulfapyridine was prolonged in those with the slow acetylator genotype. A *NAT2* 341T>C polymorphism seemed to contribute the strongest functional effect on *NAT2* activity in slow acetylators [122].

Thus, the acetylator status of a patient as determined by the *NAT2* genotype appears to be an important determinant of the risk for SSZ toxicity based on the limited data published so far (Table 29.4). While more studies and data are clearly needed, the data from the studies discussed suggest that prospective screening of patients for the *NAT2* genotype prior to initiation of SSZ may be a useful tool to prevent SSZ toxicity.

29.2.5 Pharmacogenetics of Hydroxychloroquine

Antimalarials such as hydroxychloroquine (HCQ) are effective treatments for RA and SLE. Pharmacogenetic data are available for this drug in SLE but not in RA. HCQ is rapidly

TABLE 29.4 Pharmacogenetics of Sulfasalazine

Polymorphism	Effect	Associated Biochemical Changes	Clinical Effects	Reference
*NAT2*5A*	Decreased activity of NAT2 enzyme leading to slow acetylation (slow acetylator)	Increased concentrations of SSZ intermediates due to slow acetylation	Agranulocytosis	[120]
			Fever, rash	[121]
*NAT2*5B*	As above	As above	As above	[120,121]
*NAT2*5C*	As above	As above	As above	[120–122]
*NAT2*6*	As above	As above	As above	[120,121]
*NAT2*7*	As above	As above	As above	[120,121]

Source: Modified from Davila and Ranganathan et al. [34].

absorbed from the gut after oral administration and is metabolized in the liver by cytochrome P450 (CYP) enzymes to its active metabolite, N-desethylhydroxychloroquine [34] (Figure 29.5). N-desethylhydroxychloroquine is a weak base that accumulates in acidic vesicles, such as cellular endosomes and lysosomes, increasing the pH of these compartments [123]. Consequently, its presence inhibits the action of acidic proteases involved in multiple cellular functions, particularly intracellular protein processing.

Subsequent downstream effects include interference in antigen presentation and production of proinflammatory and anti-inflammatory cytokines such as tumor necrosis factor (TNF) and interleukin (IL)-10 [124,125] Antimalarials may also have negative effects on the production of a variety of proinflammatory cytokines besides TNF, such as IL-1, IL-6, and IFNγ, although such effects have been shown only *in vitro* [126]. The lysomotropic properties of antimalarials may also inhibit the interaction between intracellular toll-like receptors (TLRs), especially TLR 9, and extracellular nucleic acid motifs in SLE [127].

29.2.5.1 CYP

Low serum levels of HCQ have been associated with high disease activity and higher risk of disease exacerbation in SLE. Variants in the genes encoding the CYP enzymes that metabolize hydroxychloroquine may influence the efficacy or toxicity of the drug. CYP2C8, CYP3A4, and CYP2D6 are the major enzymes responsible for chloroquine metabolism in human liver microsomes, [128,129], with CYP2C8 being the most important. The metabolism of the antimalarial amodiaquine (structurally related to chloroquine) appears to be impaired in individuals with *CYP2C8* polymorphisms [130].

A few studies have looked at the influence of CYP polymorphisms on HCQ metabolism. Wahie et al studied a cohort of 200 patients with discoid lupus erythematosus (DLE) treated with HCQ. They found no association between *CYP2C8* and *2D6* variants and response to HCQ (OR for *CYP2C8* 1.55 (0.82–2.96); for *CYP2D6*, 1.20 (0.67–2.18)). However, there was a trend for *CYP2C8* variants to be associated with a greater response to HCQ (although not statistically significant), as these variants may lead to a slower metabolism of antimalarials [131].

29.2.5.2 TNF/IL-10

Polymorphisms in certain cytokine genes such as *TNF* and *IL-10* have been studied for their ability to predict response to HCQ in patients with SLE. An *IL10* –1082A>G polymorphism, along with two others (*IL10*–819C>T and –592C>A), all in the *IL10* promoter region, influence basal and induced IL-10 production, with the GCC/GCC haplotype conferring the highest level of IL-10 production [132–134]. The *TNF*-308A allele is associated with increased transcription and synthesis and high serum levels of TNF as compared to the *TNF*-308G allele [135,136]. (This variant is discussed further in Section 29.3.) In a case-control study from Spain, serum TNF levels were quantified in 171 SLE patients and 215 healthy controls by ELISA techniques,

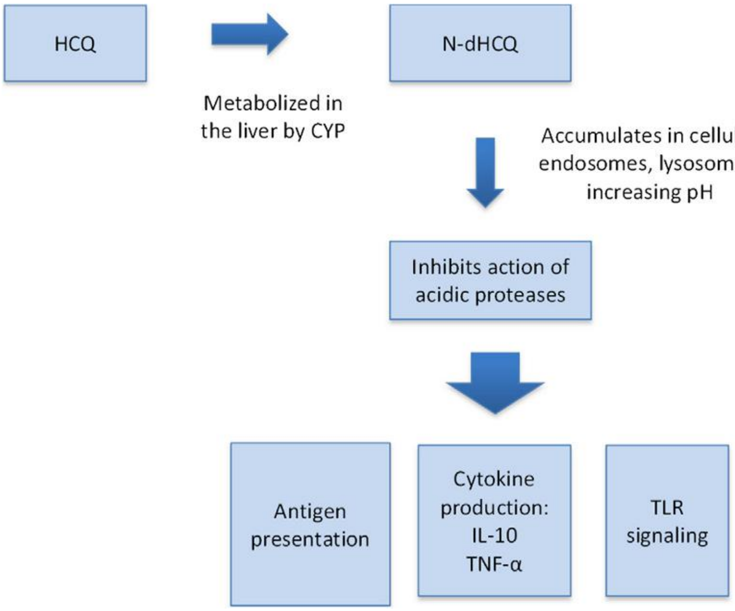

HCQ
N-dHCQ
Metabolized in the liver by CYP
Accumulates in cellular endosomes, lysosomes, increasing pH
Inhibits action of acidic proteases
Antigen presentation
Cytokine production: IL-10 TNF-α
TLR signaling

FIGURE 29.5 Hydroxychloroquine metabolism. HCQ—Hydroxychloroquine; CYP—Cytochrome P450 enzymes; N-dHCQ - N-desethylhydroxychloroquine; IL-10—Interleukin 10; TNF-α– Tumor necrosis factor alpha. *Source: Reproduced from Davila and Ranganathan et al.* [34].

TABLE 29.5 Pharmacogenetics of Hydroxychloroquine

Gene, Product, and Role	Variant(s)	Cellular Effects of Variant(s)	Clinical Effects (Reference)
IL-10, IL-10, proinflammatory cytokine	1082 A>G 819 C>T 592 C>A	Alters *IL10* promoter region, influencing basal and induced IL-10 production	Associated with increased efficacy [137]
TNF, TNF, proinflammatory cytokine	308 A>G	Alters *TNF* promoter activity and serum TNF levels	Associated with increased efficacy [137]
CYP, CYP Enzyme involved in HCQ metabolism in the liver	2C8 and 2D6	Impairs metabolism of hydroxychloroquine	Not associated with response [131]

Source: Modified from Davila and Ranganathan et al. [34].

together with genotyping for the *IL10* –1082A>G and *TNF* –308A>G promoter gene polymorphisms in 192 patients with SLE and 343 matched healthy controls [137]. Among the 192 patients who were genotyped for SNPs in both cytokine genes, 40 (20.8%) were good responders to antimalarial therapy whereas 74 (38.5%) required the addition of corticosteroids or other immunosuppressive drugs for an optimal response. Patients with SLE had higher serum TNF levels compared with healthy controls, but patients with SLE treated with antimalarials had lower serum TNF levels.

In this study, there was a significant association between carriage of the high TNF producer genotype (AA/GA) and a good response to antimalarial drugs (OR 2.25; 95% CI 1.11–4.58; $p = 0.024$), whereas the *IL-10* genotype by itself did not show any significant association. However, patients who carried both genotypes, which translated into low IL-10 and high TNF production (that is, *IL10* 1082 AA/AG and *TNF*-308 AA/GA), had the best response to HCQ (OR 3.13; 95% CI 1.41–6.92; $p = 0.005$) [137].

Thus, despite the widespread use of HCQ in RA and SLE, pharmacogenetic data that could guide the optimum use of HCQ are currently sparse and available only for SLE (Table 29.5). Further studies of possibly similar SNPs in RA and larger cohorts of SLE patients are needed to replicate and confirm the findings described.

29.3 PHARMACOGENETICS OF BIOLOGIC DMARDS

29.3.1 Tumor Necrosis Factor Antagonists

Over the past decade, tumor necrosis factor alpha (TNF-α) antagonists have become the mainstay of treatment for severe RA. Anti-TNF agents are very effective in slowing down and even halting the radiographic progression of the disease; this benefit is evident even in patients who may not have a complete clinical response [138,139]. However, anti-TNF agents fail in about 20–40% of RA patients, and while lack of response is drug-specific in some patients, it may be class-specific (anti-TNF class) in others [140,141].

Currently approved anti-TNF agents include etanercept (ETN), infliximab (INF), adalimumab (ADA), golimumab (GOL), and certolizumab pegol (CZP). INF, ADA, and GOL are all monoclonal antibodies to TNF-α; INF is a chimeric mouse protein antibody, while ADA and GOL are fully humanized antibodies. ETN, the first anti-TNF agent to be used in RA, is a fusion protein of the Fc portion of human IgG1 attached to two identical chains of the p75 portion of the TNF receptor; it binds to soluble TNF-α. CZP is a human anti-TNF Fab fragment that is conjugated with polyethylene glycol to increase its half-life.

Anti-TNF agents are expensive and have significant side effects, including an increased risk of infections. At present, there are no means to predict which individual patient may benefit from a certain drug. Over the past few years, several studies have attempted to define pharmacogenetic markers to predict RA response to anti-TNF therapy. Various genes in the TNF pathway have been evaluated using a candidate gene approach. Polymorphisms in genes for TNF-α, TNF receptors, signaling pathways downstream of the TNF receptors, Fc receptor genes, and risk variants for RA have been studied. We review these studies as well as genome-wide association studies (GWASs) of anti-TNF therapy in RA in the following sections (Table 29.6).

29.3.1.1 TNF Gene Polymorphisms

The TNF family consists of the cytokines TNF-α and lymphotoxins A (LTA) and B (LTB). The TNF locus, consisting of genes for TNF, LTA, and LTB, is on chromosome 6 and in close proximity to the HLA B and MHC III DR regions (Figure 29.6).

The promoter region polymorphism in the *TNF* gene, *TNF*-308A>G is probably the most significant and best studied of the *TNF* gene polymorphisms. The *TNF*-308A

TABLE 29.6 Pharmacogenetics of TNF Antagonists

Genes/Polymorphisms	Postulated Effect of Gene/Polymorphism	Clinical Effects	Reference
TNF promoter-308A> G	May increase transcription of TNF-α gene	Increased response to INF	[135,144]
	May increase circulating TNF-α levels	No effect on response to INF	[142,143]
TNF promoter-238G>A	May increase transcription of TNF-α gene	No effect on response to ETN	[145]
TNF +489 G/G	Intronic polymorphism—function unknown	No effect on response to ETN	[146]
TNFRSF1A-609, -580, -383	May affect ligand binding	No effect on response to ETN	[147]
TNFRSF1B 196 T>G	May affect receptor shedding and ligand binding; may increase synthesis of IL-6	Increased response to INF, ETN	[146]
		No effect on response to ETN	[147]
TNFRSF1B 676T>G	May affect membrane receptor shedding	Increased response to anti-TNF therapy	[148]
TNF microsatellites a, b, c, d, e	May influence production of TNF by PBMC, linked to TNF-308 SNP, increased susceptibility to RA	Specific *TNFa/b* haplotype associated with response to INF	[149]
		No effect on response to ETN	[147]
HLA DR, DQ alleles	May increase susceptibility to and severity of RA, close proximity to TNF locus	No effect on response to INF	[149]
		Specific individual *HLA DRB1* alleles and haplotype markers of increased response to ETN	[147]
MHC class I chain-related gene A transmembrane polymorphism	As above	No effect on response to INF	[149]
HLA microsatellites *BAT2, D6S273, D6S2223*	Haplotype may carry "response gene"	*BAT2-D6S273* haplotype associated with increased response to INF	[149]
FcγRIIIA 158FF	Low affinity for IgG, affects antibody clearance	Increased response to anti-TNF therapy	[150–152]
FCγRIIA-131RR	Low affinity for IgG, affects antibody clearance	Increased response to IFN	[151]
PTPRC rs10919563	RA susceptibility marker	Associated with good response to anti-TNF therapy	[153,154]
Polymorphisms in *MAP* kinase pathways *MAP3K1* rs96844, *MAP3K14* rs4792847, *MAP2K6* rs11656130	Affect signaling pathways in RA	Associated with good response to anti-TNF therapy	[155,156]
AFF3 SNPs rs10865035, rs1160542	RA susceptibility markers	Associated with good response to anti-TNF therapy	[157]
CD226 SNP rs763361	T and NK cell cytotoxicity	Associated with good response to anti-TNF therapy	[157]
IL-10 promoter microsatellite polymorphism R3	Associated with IL-10 secretion	Good response to ETN	[158]
-1087G>A in *IL-10* in combination with -308TNF	Associated with IL-10 secretion	Good response to ETN	[159]
IL-6 -174GG	Associated with IL-6 levels	Good response to ETN	[160]

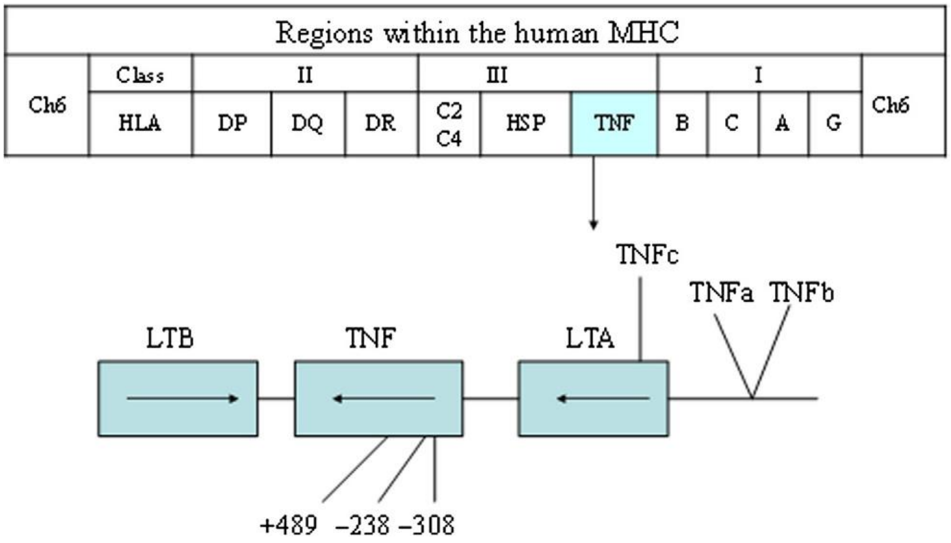


FIGURE 29.6 **TNF locus with some of the polymorphic sites known within the TNF locus.** C2, C4—Complement C2, C4; Ch—Chromosome; HLA—Human leukocyte antigen; HSP—Heat shock protein; LTA—lymphotoxin alpha; LTB—lymphotoxin beta; MHC—Major histocompatibility complex; TNF—Tumor necrosis factor. *Source: Reproduced from Davila and Ranganathan et al.* [34].

allele is associated with increased transcription and higher levels of TNF-α as compared to the *TNF*-308G allele. In an early study of 59 RA patients, those without the A allele had a better response to INF by disease activity score (DAS), than those with the A allele (42%) ($p = 0.0009$) [135]. The poor response associated with the A allele is likely related to higher TNF-α levels in patients carrying it. Cuchacovich et al. [136] genotyped 132 patients with RA for the *TNF*-308 promoter polymorphism. Of these, ten patients with *TNF*-308 G/A and 10 with *TNF* -308 G/G were given INF. INF therapy caused elevation in TNF-α in both groups, which had a similar response to treatment. Although both groups showed a similar response, the increase in TNF-α levels correlated with the ACR50 response, another standardized measure of disease activity in RA, only in patients with the G/A polymorphism ($p < 0.03$). The researchers postulated that the *TNF*-308 polymorphism influences response to INF by its effects on circulating TNF-α levels.

Subsequently, several studies looked at the effects of *TNF*-308 on response to anti-TNF therapy in RA, arriving at mixed results. Three recent meta-analyses analyzed these studies but also yielded mixed results. Two meta-analyses by Lee et al. and Pavy et al. did not find any significant association between the *TNF*-308 polymorphism and response to anti-TNF therapy in RA [142,143]. Another, by Zeng et al., which evaluated 15 studies with a total of 2127 patients with RA, did find the *TNF*-308G allele to be associated with an improved response to anti-TNF agents [144].

Another SNP in the promoter region, *TNF*-238G>A (rs361525), also has possible effects on TNF-α production. In a study by Maxwell et al. [146], the *TNF*-238G>A genotype predicted a poor response to INF but not to ETN ($p = 0.028$, $n = 40$) [145]. Other *TNF* polymorphisms of possible significance include the promoter region SNP −857C>T and the intronic SNP +489, which are known to be associated with severe RA. However, no conclusive data exist regarding their association with response to anti-TNF therapy.

29.3.1.2 TNF Receptor Polymorphisms

Soluble TNF-α binds to two transmembrane receptors: the p55 receptor, also known as TNF receptor type 1 (TNFRSF1A); and the p75 receptor, also known as TNF receptor type 2 (TNFRSF1B). Production and regulation of soluble TNFRs modulate TNF-α activity in RA joints. Therefore, polymorphisms in the TNF-α receptors have been of interest in determining response to anti-TNF therapy. The *TNFRSF1B* gene is located on chromosome 1 and has 10 exons and 9 introns [161]. The SNP *TNFRSF1B* 196T>G in exon 6 leads to an amino acid substitution of methionine to arginine [162]. This polymorphism was studied in 175 RA patients, 66 of whom were treated with either ETN or INF, and their response to treatment was assessed using the DAS. Of those receiving TNF antagonist therapy, 38 had the TT, 22 had the TG, and 6 had the GG genotypes. The GG genotype was more prevalent in patients with severe RA (6.4%) as compared to the mild or moderate group (3.1%). The TT genotype was predictive of a better response to therapy over 24 weeks compared to the TG or GG genotypes. The greatest difference in response based on the genotype was seen at 12 weeks (OR 5.1; CI 1.3–19.96; $p = 0.03$) [146].

Another polymorphism in exon 6 of the *TNFRSF1B* gene, 676T>G, also results in an amino acid change from methionine to arginine. In a different study, the 676TT genotype predicted a better response to anti-TNF therapy as compared to the 676TG genotype at 3 months (OR 3.78; 95% CI 1.07–13.31) and 12 months (OR 4.30; 95% CI 1.16–15.99) [148]. In some studies, this association was only seen in patients who were anti-CCP-positive [148,163]. A larger study by Criswell et al. genotyped 457 RA patients for the 676T>G polymorphism, but could not confirm any association with anti-TNF therapy [147].

29.3.1.3 MHC Gene Polymorphisms, TNF, and MHC Microsatellites

The "*TNF locus*," *HLA B*, and *HLA DR* genes (*MHC* genes) are all arranged in tandem on chromosome 6. As described earlier, there is a strong link between specific HLA DRB1 alleles (also called the shared epitope alleles) and susceptibility to RA and its severity [164]. This makes it likely that *MHC* gene polymorphisms may affect the response to anti-TNF agents. Microsatellite haplotypes were previously associated with susceptibility to RA and certain *TNF* promoter region SNPs [165,166]. In a study by Martinez et al., 78 RA patients on INF therapy were genotyped for *HLA-DRB1*, *HLA-DQA1*, *HLA-DQB1*, *MHC* class I chain-related gene A (*MICA*) transmembrane polymorphism alleles, microsatellites *TNFa-e*, *D6S273*, HLA-B associated transcript 2 (*BATS2*), and *D6S2223*. Some interesting associations were observed between certain microsatellite haplotypes and response. Among the microsatellite haplotypes, the *D6S273_4/BAT2_2* pair was a marker of the INF responder group, both among patients and when compared to controls (46% versus 11% in nonresponders, $p = 0.001$; 46% in responders versus 17% in controls, $p = 0.00002$), indicating that this microsatellite pair may occur on the haplotype that carries the "response gene," or each microsatellite allele may be a marker of a "response gene" in proximity. The frequency of one of the *TNFa/b* haplotypes was increased in responders compared to nonresponders (41% versus 16% in nonresponders, $p = 0.01$). None of the *MHC* alleles influenced response to INF. Thus some microsatellite haplotypes were associated with response to INF in this study; single alleles did not reveal similar associations [149].

In a second study, patients were genotyped for the "shared epitope" specific *HLA-DRB1* alleles. Genotyping was carried out on 457 patients with active RA of ≤3 years duration, who were categorized as having 0, 1, or 2 copies of the SE. Patients were treated with MTX or ETN for one year, and response to therapy was measured by ACR50. SNPs at positions −308, −238, and +488 of the *TNF* gene; +249, +365, and +720 of the *LTA* gene were also examined. (These 6 *LTA-TNF* SNPs mark haplotypes spanning the "TNF locus" region). Five *TNF* microsatellites (*TNF a–e*), SNPs in *TNFRSF1A* at positions −609, −580, and −383, and the 196T/G polymorphism in *TNFRSF1B* were also examined.

As the Fc receptor (FcR) pathway appears important in the degradation of ETN-TNF complexes, three *FcR* polymorphisms were also examined in this study. Response to treatment was associated with a higher number of *HLA-DRB1* (SE) alleles. Patients with 2 copies of the SE alleles had a better response to ETN compared to those with 1 copy or none (OR 4.3; 95% CI 1.8–10.3). Haplotypes defined by the 6 *LTA—TNF* SNPs and *DRB1* alleles were deduced for the 16 most common *DRB1* alleles in a subset of 224 Caucasian patients. Among 448 haplotypes thus examined, two, *HLA-DRB1* *0101-GGGAGG and *HLA-DRB1* *0404-GGAAGG, strongly correlated with response (76 and 61% ACR50 response at 12 months, respectively). The number of copies of *HLA-DRB1* SE alleles inherited and specific haplotypes spanning the *HLA-DRB1* region and SNPs in the *LTA–TNF* region may therefore be associated with response to ETN, at least in Caucasians [147].

29.3.1.4 RA Risk Variants

Several risk alleles have been identified that are associated with increased susceptibility to RA [167]. Given this association, these alleles may also influence response to treatment. Some data suggest that the RA risk variant gene tyrosine phosphatase receptor type C (*PTPRC*) has a strong association with response to anti-TNF therapy. The *PTPRC* gene encodes a CD45 antigen. In a study by Cui et al. [153], a cohort of 1283 RA patients on TNF inhibitors were genotyped for SNPs in *PTPRC*. Patients with the rs10919563 SNP in the *PTPRC* gene had a better EULAR response as compared to patients without the polymorphism ($p = 0.0001$). Another large cohort of 1115 RA patients from the United Kingdom were tested for 29 SNPs known to be RA susceptibility variants. The rs10919563 SNP in the *PTPRC* gene was again associated with improved response to anti-TNF therapy (OR 0.62; 95% CI 0.40–0.95; $p = 0.03$) [154]. Despite the relatively strong association found in these studies, a subsequent study by Krintel et al. was unable to replicate these findings [168].

Mitogen-activated protein kinases (MAPKs), which are involved in several signaling pathways in RA, are also known RA risk variants. Studies evaluating polymorphisms in the MAPK pathways have yielded mixed results. In a UK cohort of 642 patients, the SNP rs96844 in *MAP3K1* and the SNP rs4792847 in *MAP3K14* were found to be nominally associated with response to anti-TNF therapy ($p < 0.05$). However, these associations could not be confirmed in a validation cohort of 428 patients with RA in the same study [155]. Coulthard et al. [156] studied polymorphisms in 11 genes in the MAPK pathway in a cohort of 1102 RA patients receiving TNF inhibitors. Seven polymorphisms were identified in five genes in the MAPK signaling pathways that were associated with an improved DAS28 response to IFN and ADA. However, no association was found with response to ETN. One polymorphism, rs11656130 in *MAP2K6*, was associated with a good EULAR response but not an improved response as measured by DAS28.

Two more SNPs in *AFF3* were identified as markers of response to anti-TNF therapy by Tan et al. [157] in 1012 RA patients. The G allele of two SNPs, rs10865035 and rs1160542, in *AFF3* were associated with response to anti-TNF therapy. The *AFF3* gene, which encodes nuclear transcription activators in lymphoid tissue, is a known RA susceptibility marker. Another SNP rs763361 in the *CD226* gene results in a substitution of serine for glycine. This SNP was also associated with response to anti-TNF therapy [169]. CD226, a membrane protein on the surface of hematopoietic cells, is involved in T and NK cell cytotoxicity and has been described as a RA risk variant [170].

29.3.1.5 Cytokines, Toll-Like Receptors, and Signaling Pathways

Multiple cytokines are implicated in RA pathogenesis. Polymorphisms in the interleukin (*IL)-1β*, IL-1 receptor antagonist (*ILRA*), *IL-6*, *IL-10*, and tumor growth factor (*TGF*) β genes have been studied in relation to anti-TNF therapy. So far, no polymorphisms in individual cytokine genes have been associated with response [171]. According to Padyukov et al. [159], a −1087G>A SNP in the *IL10* gene, in combination with the −308 SNP in the *TNF* gene, were associated with a good response to ETN. Another combination, the A2 allele in intron 2 of the *ILRA* gene and a rare +915C allele in codon 25 of the *TGFB1* gene, was associated with a poor response to ETN therapy. The +915C allele was found significantly more often in combination with the A2 allele in the *ILRA* gene in nonresponders to ETN ($p<0.05$) [159]. Schotte et al. [158] reported that the *IL-10* microsatellite polymorphisms—the R3 allele and the R3-G9 haplotype in the promoter region—were associated with a good response, whereas the G13 allele and the haplotype R2-G13 were associated with moderate to no response [158].

Jancic et al. [160] genotyped 77 RA patients for the 174G>C*IL-6* gene polymorphism, which is known to influence IL-6 levels. These patients were treated for 12 months with ETN. More patients with the *IL-6*-174GG (95.7%) genotype had an improvement in disease activity by DAS as compared to those with the GC (75.6%) or CC (44.4%) genotypes ($p=0.006$). Potter et al. [172] examined SNPs in the NFkB (or nuclear factor kappa B, a transcription factor) and TLR (toll-like receptor) pathways. Among 187 SNPs genotyped in 909 RA patients, 12 in 9 genes showed a nominal association with response to anti-TNF therapy.

29.3.1.6 Fcγ Receptor Variants

Most anti-TNF agents are antibodies and therefore have a Fc component to them (except CZB, which is a pegylated Fab fragment). The Fc component attaches to the Fc gamma receptor (FcγR) on various cells and serves as a mechanism for antibody clearance. Polymorphisms in the *FcγR* gene affect the avidity and strength of Fc–FcγR binding and thereby may potentially influence the clearance and thus the efficacy of TNF antagonist drugs. For example, in the *FcγRIIIA* receptor polymorphism F158V, the FF genotype has a lower affinity for IgG1. In a small study of 30 patients with RA, the 158 FF variant was found in greater frequency among patients who had a very good response to anti-TNF agents, presumably because of lower drug clearance rates [150]. In another study of 98 RA patients treated with IFN, both ACR50 and EULAR responses at week 6 were better in those who were homozygous for the low-affinity *FCγRIIIA* 158FF genotype. Another affinity variant, *FCγRIIA* 131RR, was a marker of better ACR20 responses at week 30 of INF treatment in the same study [151]. These findings were replicated in a Japanese study with 33 patients, [152], but they could not be replicated in a subsequent larger Swedish study of 282 patients treated with INF or ETN [173].

29.3.1.7 GWASs

Genome-wide association studies (GWAS) represent another methodology used to evaluate the association of genetic polymorphisms with response to anti-TNF therapy. While they are a powerful way to look at a multitude of SNPs, higher numbers of patients are required to achieve significant results. Also, while the SNPs may not be significant in themselves, they may point to a genetic locus of significance. Liu et al. [174] reported the first GWAS in this area using 89 patients; there was an association of 16 SNPs in nine genes with response to anti-TNF therapy. These findings were not replicated in larger subsequent studies.

In a larger cohort of 566 patients, Plant et al. [175] found seven loci of significance by multistage GWAS. The first SNP, rs17301249, mapped to the eyes-absent homolog 4 (*EYA4*) gene on chromosome 6, and showed a significant association with response. *EYA4* is a co-transcription factor for interferon β and CXCL10. A second SNP, rs1532269, mapped to the PDZ domain–containing protein 2 (*PDZD2*) gene, which is associated with insulin secretion. (*PDZ* is an acronym combining the first letters of three proteins—post-synaptic density protein, *Drosophila* disc large-tumor suppressor, and *Zonula occludens*-1 protein—which were first discovered to share a domain). The significance of this locus is unclear. The other five SNPs mapped to intergenic regions on chromosomes 1, 4, 11, and 12.

Another GWAS by Krintel et al. [167] analyzed 486,450 SNPs in 196 Danish patients with moderate to severe RA on IFN, ETN, or ADA therapy. The findings of the earlier GWASs by Lui et al. and Plant et al. could not be confirmed in this study, and despite the multitude of SNPs studied none achieved significance.

A recent GWAS by Umicevic Mirkov et al. [176] analyzed 2,557,253 SNPs in 882 Dutch patients from the Dutch Rheumatoid Arthritis Monitoring (DREAM) registry. Three, (rs1568885, rs1813443, and rs4411591) showed directional consistency, and eight genetic loci were suggestive of association. However, no individual SNP achieved significant association with response to anti-TNF treatment. Also, none of the associations found in earlier studies could be confirmed.

Cui et al. recently conducted a GWAS meta-analysis with 2706 RA patients from 13 different cohorts. More than 2 million SNPs were evaluated for association with response to anti-TNF therapy. The rs6427528 SNP was associated with a better response by DAS scores in patients treated with ETN; this may be mediated by higher CD84 gene expression in peripheral blood mononuclear cells in patients carrying this SNP. CD84 is a cell-surface receptor found on various immune cells and is thought to play a role in signaling T-cell activation and IFN-γ secretion [177,178]. The variant rs6427528 was not associated with response to therapy with either INF or ADA [179]. Acosta-Colman et al. [180] described the SNP rs3794271 in intron 4 of the *SLCO1C1* gene in a cohort of 315 RA patients. This SNP was associated with a good EULAR response to anti-TNF therapy. *SLCO1C1* is a cell membrane transporter important in drug metabolism. The *SLCO1C1* locus is near the *PDE3A* gene (which codes for a phosphodiesterase). The aforementioned GWAS study by Krintel et al. had described an intergenic SNP, rs11045392, located between the 5′ end of *SLCO1C1* and the 3′ end of the *PDE3A* gene as a putative marker of response to anti-TNF therapy. The researchers speculated that these two SNPs, rs3794271 and rs11045392, may be in linkage disequilibrium and therefore that the *PDE3A-SLCO1C1* locus may be an important determinant of response to anti-TNF therapy [180].

29.3.2 Rituximab

Rituximab is an anti-CD20 chimeric antibody used in RA treatment. Several studies have looked at candidate genes for pharmacogenetic associations that affect response to rituximab therapy. Some of these genes had already been evaluated in patients with lymphoma, a condition for which rituximab has been in use for several years.

In a study of 63 RA patients, Daien et al. [181] analyzed 13 SNPs in nine genes, including *IL10, LTA*, *TGFβ1*, *TNF-α*, *TNFRSF1B*, *TRAF1-C5*, *STAT4*, *TNFAIP3*, and *PTPN22*. Some of these—rs2476601 in *PTPN22*, rs7574865 in *STAT4*, rs1081848 in *TRAF1-C5*, and rs6920220 in *TNFAIP3*—had been associated with RA susceptibility in prior studies [182,183].

In the *TGFβ1* gene, the CT genotype at SNP rs1800470 in codon 10 was associated with good treatment response to rituximab. The CT genotype, as compared to the TT genotype, was seen more often in responders (OR 1.6; 95% CI 1.2–2.3; $p = 0.002$), while the CC genotype was equally prevalent among responders and nonresponders. Another SNP, rs1800471 at codon 25 in the *TGFβ1* gene, was also associated with response to rituximab. Of patients with the GC genotype 100% were responders, while 63% with the GG genotype were responders (OR 1.6; 95% CI 1.3–1.9; $p = 0.025$) [181]. Rituximab was even more effective in patients who had both these polymorphisms—that is, SNP rs1800470 in codon 10 and SNP rs1800471 in codon 25 (OR 2.6; $p = 0.008$) [181]—which may be the result of lower TGFβ1 production in patients carrying these SNPs [184].

The B-lymphocyte–stimulating factor (BLyS) is essential for B-cell production and survival. As rituximab therapy causes depletion of B cells, it is also associated with a reflexive increase in BlyS levels that attempts to establish B-cell repopulation. Therefore, polymorphisms in BlyS may alter response to rituximab by altering B-cell repopulation after rituximab therapy. Ruyssen-Witrand et al. [185] studied the polymorphism −871C>T in the *BlyS* promoter region in 115 patients with RA. The 871CC genotype was associated with a better EULAR response to rituximab than the 871TT genotype (OR 6.9; 95% CI 1.6–29.6; $p = 0.03$), possibly mediated by a change in BlyS levels. This association of the −871C>T SNP with response to rituximab could not be validated in a subsequent study by Fabris et al. [186] in 155 Italian RA patients. In this study, four polymorphisms in the *BLyS* promoter region: −2841T>C, −2704T>C, −2701T>A, and −871C>T, in linkage disequilibrium, were examined. The haplotype TTTT was predictive of a good response to rituximab (OR 14.4; 95% CI 1.77–117.39; $p = 0.003$) in RA patients who were seropositive (positive for RF and/or CCP antibody) and had a prior poor response to anti-TNF agents. These findings were replicated in a validation cohort of 115 RA patients in the same study [186].

Because rituximab is a monoclonal antibody, its serum levels and rates of clearance are altered by Fc receptor polymorphisms. The F158V polymorphism in the *FcγRIIIA* gene (rs396991) has been associated with response to rituximab, although with variable results. In a study of 111 patients, Ruyssen-Witrand et al. [187] found that V allele carriage was associated with better response (91% of responders versus 70% of nonresponders (OR 4.6; 95% CI 1.5–13.6; $p = 0.006$). Another study of 212 RA patients reported similar findings: a good EULAR response was seen in 89.5% of patients with the VV genotype, in 66% with the VF genotype, and in 66.2% with the FF genotype ($p = 0.01$) [188]. Kastbom et al., however, found that heterozygotes carrying the 158VF genotype responded better than those homozygous for either 158VV or 158FF [189]. Sarsour et al. found no effects of the F158V polymorphism on response to either rituximab or TNF inhibitors [190].

Fabris et al. studied the −174G>C polymorphism in *IL-6* and the D358A polymorphism in IL-6-receptor alpha (*IL-6Ra*) genes for associations with rituximab response; these polymorphisms influence expression of IL6 and

IL6Ra, respectively. IL6 is a growth and maturation factor for B cells. The *IL-6*-174CC genotype (rs1800795) was associated with a poor EULAR response (39.1%) as compared to the GC/CC genotypes (18.5%) (OR 2.83; 95% CI 1.10–7.27; p = 0.031). No association was found between the D358A polymorphism in *IL-6Ra* and response to rituximab [191].

29.3.3 Tocilizumab

Tocilizumab is a humanized monoclonal IL-6 receptor antagonist that was recently approved for RA therapy. Wang et al. conducted the first GWAS to demonstrate that genetic variants are associated with response to tocilizumab. They used pooled data from six studies of tocilizumab in RA and a total of 1683 patients. In the initial analysis, 534,053 markers were studied, and 253 variants showed an association with tocilizumab therapy. Seven of these variants (rs11052877, rs4910008, rs9594987, rs10108210, rs703297, rs703505, and rs1560011) achieved significance on conformational analysis. Eight putative genetic loci of interest were identified that were not previously known to be significant. Associations of these variants with response were found to be stronger in seropositive RA patients [192].

29.4 CONCLUSIONS AND FUTURE DIRECTIONS

As seen from the studies described in this chapter, the field of pharmacogenetics in RA is rapidly evolving. Variations in genes controlling drug-metabolizing enzymes, drug receptors, and drug targets play a role in determining an individual's response to a given drug. Nonetheless, several points need to be considered before pharmacogenetics can be brought to the clinic.

In many of the studies described, the strength of the association between genotype and phenotype may not be robust, for several reasons. Most of the studies suffered from underpowering and small sample sizes, and associations observed in one study could not be replicated in another. Some were retrospective in design, which prevented an accurate estimation of drug effects, particularly adverse ones. In addition, race may strongly impact pharmacogenetics, and genotype-phenotype associations observed in one racial population may not necessarily hold true in a different racial population. Our study examining the frequencies of SNPs in the MTX pathway in different racial groups demonstrated significant differences in the allele frequencies of several SNPs between Caucasians and African Americans with RA [193]. Another study, by Criswell et al., showed that certain *MHC/TNF* haplotypes were predictive of response to ETN in a Caucasian population; whether these associations will be valid in other populations remains in question [147].

Most drugs used in RA (e.g., MTX, TNF antagonists) influence several different cellular (and genetic) pathways. Thus, analysis of SNPs in different metabolic pathways rather than only one may be more helpful in predicting response [57]. Also, if a variant is only weakly associated with response, this may be because it occurs in tandem or in proximity to the actual marker of response. For these reasons, haplotype analyses may be better than single SNP analyses in predicting response [147,194]. Genome-wide association studies in pharmacogenetics are emerging, as described in this chapter, that may also help in this regard. Finally, the cost-effectiveness of pharmacogenetic testing needs to be examined before such testing can be brought into the clinical setting [195]. Drugs with a narrow therapeutic index and severe side effects, and for which there exists a firm association between a specific genotype and phenotype (usually toxicity), are best suited for pharmacogenetic testing in the clinic; a good example is *TPMT* testing prior to AZA use in clinical practice.

Despite these caveats, as genotyping becomes easily available and less expensive, and as major funding agencies continue to show a commitment to pharmacogenetic research—for example, the International HapMap Consortium (www.hapmap.org) and the the National Institutes of Health Pharmacogenetics Research Network (www.nigms.nih.gov/pharmacogenetics/), personalized medicine in RA will become a reachable goal in the not too distant future.

REFERENCES

[1] Joseph A, Brasington R, Kahl L, Ranganathan P, Cheng TP, Atkinson J. Immunologic rheumatic disorders. J Allergy Clin Immunol 2010;125(2 Suppl 2):S204–15.

[2] Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. Curr Opin Rheumatol 2010;22(3):293–8.

[3] Turesson C, Schaid DJ, Weyand CM, Jacobsson LT, Goronzy JJ, Petersson IF, et al. The impact of HLA-DRB1 genes on extra-articular disease manifestations in rheumatoid arthritis. Arthritis Res Ther 2005;7(6):R1386–93.

[4] Mattey DL, Dawes PT, Gonzalez-Gay MA, Garcia-Porrua C, Thomson W, Hajeer AH, et al. HLA-DRB1 alleles encoding an aspartic acid at position 70 protect against development of rheumatoid arthritis. J Rheumatol 2001;28(2):232–9.

[5] Guthrie KA, Tishkevich NR, Nelson JL. Non-inherited maternal human leukocyte antigen alleles in susceptibility to familial rheumatoid arthritis. Ann Rheum Dis 2009;68(1):107–9.

[6] Rak JM, Maestroni L, Balandraud N, Guis S, Boudinet H, Guzian MC, et al. Transfer of the shared epitope through microchimerism in women with rheumatoid arthritis. Arthritis Rheum 2009;60(1):73–80.

[7] Liao KP, Alfredsson L, Karlson EW. Environmental influences on risk for rheumatoid arthritis. Curr Opin Rheumatol 2009;21(3):279–83.

[8] Criswell LA, Merlino LA, Cerhan JR, Mikuls TR, Mudano AS, Burma M, et al. Cigarette smoking and the risk of rheumatoid arthritis among postmenopausal women: results from the Iowa women's health study. Am J Med 2002;112(6):465–71.

[9] Lee YH, Rho YH, Choi SJ, Ji JD, Song GG. PADI4 polymorphisms and rheumatoid arthritis susceptibility: a meta-analysis. Rheumatol Int 2007;27(9):827–33.

[10] Lee DM, Weinblatt ME. Rheumatoid arthritis. Lancet 2001;358(9285):903–11.

[11] Choy EH, Panayi GS. Cytokine pathways and joint inflammation in rheumatoid arthritis. N Engl J Med 2001;344(12):907–16.

[12] Edwards JC, Cambridge G. Sustained improvement in rheumatoid arthritis following a protocol designed to deplete B lymphocytes. Rheumatology (Oxford) 2001;40(2):205–11.

[13] Hukkanen M, Hughes FJ, Buttery LD, Gross SS, Evans TJ, Seddon S, et al. Cytokine-stimulated expression of inducible nitric oxide synthase by mouse, rat, and human osteoblast-like cells and its functional role in osteoblast metabolic activity. Endocrinology 1995;136(12):5445–53.

[14] Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham 3rd CO, et al. 2010 Rheumatoid arthritis classification criteria: an American college of rheumatology/European league against rheumatism collaborative initiative. Arthritis Rheum 2010;62(9):2569–81.

[15] Bakker MF, Jacobs JW, Welsing PM, Vreugdenhil SA, van Booma-Frankfort C, Linn-Rasker SP, et al. Early clinical response to treatment predicts 5-year outcome in RA patients: follow-up results from the CAMERA study. Ann Rheum Dis 2011;70(6):1099–103.

[16] White D, Pahau H, Duggan E, Paul S, Thomas R. Trajectory of intensive treat-to-target disease modifying drug regimen in an observational study of an early rheumatoid arthritis cohort. BMJ Open 2013;3(7). pii: e003083

[17] Moreland LW, O'Dell JR, Paulus HE, Curtis JR, Bathon JM, St Clair EW, et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of early aggressive rheumatoid arthritis trial. Arthritis Rheum 2012;64(9):2824–35.

[18] van der Heijde D, Klareskog L, Rodriguez-Valverde V, Codreanu C, Bolosiu H, Melo-Gomes J, et al. Comparison of etanercept and methotrexate, alone and combined, in the treatment of rheumatoid arthritis: two-year clinical and radiographic results from the TEMPO study, a double-blind, randomized trial. Arthritis Rheum 2006;54(4):1063–74.

[19] Chung WS, Lin CL, Peng CL, Chen YF, Lu CC, Sung FC, et al. Rheumatoid arthritis and risk of acute myocardial infarction-a nationwide retrospective cohort study. Int J Cardiol 2013;168(5):4750–4.

[20] Weinblatt ME, Coblyn JS, Fox DA, Fraser PA, Holdsworth DE, Glass DN, et al. Efficacy of low-dose methotrexate in rheumatoid arthritis. N Engl J Med 1985;312(13):818–22.

[21] Kremer JM, Phelps CT. Long-term prospective study of the use of methotrexate in the treatment of rheumatoid arthritis. Update after a mean of 90 months. Arthritis Rheum 1992;35(2):138–45.

[22] Kremer JM. Rational use of new and existing disease-modifying agents in rheumatoid arthritis. Ann Intern Med 2001;134(8):695–706.

[23] Strand V, Cohen S, Schiff M, Weaver A, Fleischmann R, Cannon G, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. Leflunomide rheumatoid arthritis investigators group. Arch Intern Med 1999;159(21):2542–50.

[24] Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, et al. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. [erratum appears in N Engl J Med 2001;344(3):240]. N Engl J Med 2000;343(22):1586–93.

[25] Hooijberg JH, Broxterman HJ, Kool M, Assaraf YG, Peters GJ, Noordhuis P, et al. Antifolate resistance mediated by the multidrug resistance proteins MRP1 and MRP2. Cancer Res 1999;59(11):2532–5.

[26] Longo GS, Gorlick R, Tong WP, Lin S, Steinherz P, Bertino JR. Gamma-glutamyl hydrolase and folylpolyglutamate synthetase activities predict polyglutamylation of methotrexate in acute leukemias. Oncol Res 1997;9(5):259–63.

[27] Barredo JC, Synold TW, Laver J, Relling MV, Pui CH, Priest DG, et al. Differences in constitutive and post-methotrexate folylpolyglutamate synthetase activity in B-lineage and T-lineage leukemia. Blood 1994;84(2):564–9.

[28] Galpin AJ, Schuetz JD, Masson E, Yanishevski Y, Synold TW, Barredo JC, et al. Differences in folylpolyglutamate synthetase and dihydrofolate reductase expression in human B-lineage versus T-lineage leukemic lymphoblasts: mechanisms for lineage differences in methotrexate polyglutamylation and cytotoxicity. Mol Pharmacol 1997;52(1):155–63.

[29] Rots MG, Pieters R, Peters GJ, Noordhuis P, van Zantwijk CH, Kaspers GJ, et al. Role of folylpolyglutamate synthetase and folylpolyglutamate hydrolase in methotrexate accumulation and polyglutamylation in childhood leukemia. Blood 1999;93(5):1677–83.

[30] Galivan J. Evidence for the cytotoxic activity of polyglutamate derivatives of methotrexate. Mol Pharmacol 1980;17(1):105–10.

[31] van Ede AE, Laan RF, Blom HJ, De Abreu RA, van de Putte LB. Methotrexate in rheumatoid arthritis: an update with focus on mechanisms involved in toxicity. Arthritis Rheum 1998;27(5):277–92.

[32] Szeto DW, Cheng YC, Rosowsky A, Yu CS, Modest EJ, Piper JR, et al. Human thymidylate synthetase–III. Effects of methotrexate and folate analogs. Biochem Pharmacol 1979;28(17):2633–7.

[33] Chan ES, Cronstein BN. Molecular action of methotrexate in inflammatory diseases. Arthritis Res 2002;4(4):266–73.

[34] Davila L, Ranganathan P. Pharmacogenetics: implications for therapy in rheumatic diseases. Nat Rev Rheumatol 2011;7(9):537–50.

[35] Hayashi H, Fujimaki C, Daimon T, Tsuboi S, Matsuyama T, Itoh K. Genetic polymorphisms in folate pathway enzymes as a possible marker for predicting the outcome of methotrexate therapy in Japanese patients with rheumatoid arthritis. J Clin Pharm Ther 2009;34(3):355–61.

[36] Dervieux T, Furst D, Lein DO, Capps R, Smith K, Walsh M, et al. Polyglutamation of methotrexate with common polymorphisms in reduced folate carrier, aminoimidazole carboxamide ribonucleotide transformylase, and thymidylate synthase are associated with methotrexate effects in rheumatoid arthritis. Arthritis Rheum 2004;50(9):2766–74.

[37] Pawlik A, Herczynska M, Kurzawski M, Safranow K, Dziedziejko V, Drozdzik M. The effect of exon (19C>A) dihydroorotate dehydrogenase gene polymorphism on rheumatoid arthritis treatment with leflunomide. Pharmacogenomics 2009;10(2):303–9.

[38] Haagsma CJ, Blom HJ, van Riel PL, Van't Hof MA, Giesendorf BA, van Oppenraaij-Emmerzaal D, et al. Influence of sulphasalazine, methotrexate, and the combination of both on plasma homocysteine concentrations in patients with rheumatoid arthritis. Ann Rheum Dis 1999;58(2):79–84.

[39] van Ede AE, Laan RF, Blom HJ, Huizinga TW, Haagsma CJ, Giesendorf BA, et al. The C677T mutation in the methylenetetrahydrofolate reductase gene: a genetic risk factor for methotrexate-related

elevation of liver enzymes in rheumatoid arthritis patients. Arthritis Rheum 2001;44(11):2525–30.

[40] Urano W, Taniguchi A, Yamanaka H, Tanaka E, Nakajima H, Matsuda Y, et al. Polymorphisms in the methylenetetrahydrofolate reductase gene were associated with both the efficacy and the toxicity of methotrexate used for the treatment of rheumatoid arthritis, as evidenced by single locus and haplotype analyses. Pharmacogenetics 2002;12(3):183–90.

[41] Wessels JA, de Vries-Bouwstra JK, Heijmans BT, Slagboom PE, Goekoop-Ruiterman YP, Allaart CF, et al. Efficacy and toxicity of methotrexate in early rheumatoid arthritis are associated with single-nucleotide polymorphisms in genes coding for folate pathway enzymes. Arthritis Rheum 2006;54(4):1087–95.

[42] Fisher MC, Cronstein BN. Metaanalysis of methylenetetrahydrofolate reductase (MTHFR) polymorphisms affecting methotrexate toxicity. J Rheumatol 2009;36(3):539–45.

[43] Berkun Y, Levartovsky D, Rubinow A, Orbach H, Aamar S, Grenader T, Abou Atta I, et al. Methotrexate related adverse effects in patients with rheumatoid arthritis are associated with the A1298C polymorphism of the MTHFR gene. Ann Rheum Dis 2004;63:1227–31.

[44] Lee YH, Song GG. Associations between the C677T and A1298C polymorphisms of MTHFR and the efficacy and toxicity of methotrexate in rheumatoid arthritis: a meta-analysis. Clin Drug Investig 2010;30(2):101–8.

[45] Grzybowska EA, Wilczynska A, Siedlecki JA. Regulatory functions of 3’UTRs. Biochem Biophys Res Commun 2001;288(2):291–5.

[46] Dervieux T, Greenstein N, Kremer J. Pharmacogenomic and metabolic biomarkers in the folate pathway and their association with methotrexate effects during dosage escalation in rheumatoid arthritis. Arthritis Rheum 2006;54(10):3095–103.

[47] Weisman MH, Furst DE, Park GS, Kremer JM, Smith KM, Wallace DJ, et al. Risk genotypes in folate-dependent enzymes and their association with methotrexate-related side effects in rheumatoid arthritis. Arthritis Rheum 2006;54(2):607–12.

[48] Grabar PB, Rojko S, Logar D, Dolzan V. Genetic determinants of methotrexate treatment in rheumatoid arthritis patients: a study of polymorphisms in the adenosine pathway. Ann Rheum Dis 2010;69(5):931–2.

[49] Kumagai K, Hiyama K, Oyama T, Maeda H, Kohno N. Polymorphisms in the thymidylate synthase and methylenetetrahydrofolate reductase genes and sensitivity to the low-dose methotrexate therapy in patients with rheumatoid arthritis. Int J Mol Med 2003;11(5):593–600.

[50] Cheng Q, Wu B, Kager L, Panetta JC, Zheng J, Pui CH, et al. A substrate specific functional polymorphism of human gamma-glutamyl hydrolase alters catalytic activity and methotrexate polyglutamate accumulation in acute lymphoblastic leukaemia cells. Pharmacogenetics 2004;14(8):557–67.

[51] Jekic B, Lukovic L, Bunjevacki V, Milic V, Novakovic I, Damnjanovic T, et al. Association of the TYMS 3G/3G genotype with poor response and GGH 354GG genotype with the bone marrow toxicity of the methotrexate in RA patients. Eur J Clin Pharmacol 2013;69(3):377–83.

[52] Rothem L, Aronheim A, Assaraf YG. Alterations in the expression of transcription factors and the reduced folate carrier as a novel mechanism of antifolate resistance in human leukemia cells. J Biol Chem 2003;278(11):8935–41.

[53] Rothem L, Stark M, Kaufman Y, Mayo L, Assaraf YG. Reduced folate carrier gene silencing in multiple antifolate-resistant tumor cell lines is due to a simultaneous loss of function of multiple transcription factors but not promoter methylation. J Biol Chem 2004;279(1):374–84.

[54] Whetstine JR, Witt TL, Matherly LH. The human reduced folate carrier gene is regulated by the AP2 and sp1 transcription factor families and a functional 61-base pair polymorphism. J Biol Chem 2002;277(46):43873–80.

[55] Dervieux T, Lein DO, Park G, Barham R, Smith K, Walsh M, et al. Single nucleotide polymorphisms in the folate/purine synthesis pathway predict methotrexate’s effects in rheumatoid arthritis [abstract]. Arthritis Rheum 2003;48(Suppl 9):S438.

[56] Chave KJ, Ryan TJ, Chmura SE, Galivan J. Identification of single nucleotide polymorphisms in the human gamma-glutamyl hydrolase gene and characterization of promoter polymorphisms. Gene 2003;319:167–75.

[57] Dervieux T, Kremer J, Lein DO, Capps R, Barham R, Meyer G, et al. Contribution of common polymorphisms in reduced folate carrier and gamma-glutamylhydrolase to methotrexate polyglutamate levels in patients with rheumatoid arthritis. Pharmacogenetics 2004;14(11):733–9.

[58] Hoffmeyer S, Burk O, von Richter O, Arnold HP, Brockmoller J, Johne A, et al. Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity *in vivo*. Proc Natl Acad Sci USA 2000;97(7):3473–8.

[59] Kim RB, Leake BF, Choo EF, Dresser GK, Kubba SV, Schwarz UI, et al. Identification of functionally variant MDR1 alleles among European Americans and African Americans. Clin Pharmacol Ther 2001;70(2):189–99.

[60] Tanabe M, Ieiri I, Nagata N, Inoue K, Ito S, Kanamori Y, et al. Expression of P-glycoprotein in human placenta: relation to genetic polymorphism of the multidrug resistance (MDR)-1 gene. J Pharmacology Exp Ther 2001;297(3):1137–43.

[61] Norris MD, De Graaf D, Haber M, Kavallaris M, Madafiglio J, Gilbert J, et al. Involvement of MDR1 P-glycoprotein in multifactorial resistance to methotrexate. Int J Cancer 1996;65(5):613–9.

[62] Mickisch GH, Merlino GT, Galski H, Gottesman MM, Pastan I. Transgenic mice that express the human multidrug-resistance gene in bone marrow enable a rapid identification of agents that reverse drug resistance. Proc Natl Acad Sci 1991;88(2):547–51.

[63] Hider SL, Morgan C, Bell E, Bruce IN. Methotrexate is not a substrate for P-glycoprotein in patients with rheumatoid arthritis. Ann Rheum Dis 2002;61(Suppl 1):199.

[64] Pawlik A, Wrzesniewska J, Fiedorowicz-Fabrycy I, Gawronska-Szklarz B. The MDR1 3435 polymorphism in patients with rheumatoid arthritis. Int J Clin Pharmacol Ther 2004;42(9):496–503.

[65] Rozen R. Molecular genetics of methylenetetrahydrofolate reductase deficiency. J Inherit Metab Dis 1996;19(5):589–94.

[66] Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, et al. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat Genet 1995;10(1):111–3.

[67] van der Put NM, Gabreels F, Stevens EM, Smeitink JA, Trijbels FJ, Eskes TK, et al. A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects? Am J Hum Genet 1998;62(5):1044–51.

[68] Kang SS, Zhou J, Wong PW, Kowalisyn J, Strokosch G. Intermediate homocysteinemia: a thermolabile variant of methylenetetrahydrofolate reductase. Am J Hum Genet 1988;43(4):414–21.

[69] Kurzawski M, Pawlik A, Safranow K, Herczynska M, Drozdzik M. 677C>T and 1298A>C MTHFR polymorphisms affect methotrexate treatment outcome in rheumatoid arthritis. Pharmacogenomics 2007;8(11):1551–9.

[70] Taniguchi A, Urano W, Tanaka E, Furihata S, Kamitsuji S, Inoue E, et al. Validation of the associations between single nucleotide polymorphisms or haplotypes and responses to disease-modifying antirheumatic drugs in patients with rheumatoid arthritis: a proposal for prospective pharmacogenomic study in clinical practice. Pharmacogenet Genomics 2007;17(6):383–90.

[71] Kim SK, Jun JB, El-Sohemy A, Bae SC. Cost-effectiveness analysis of MTHFR polymorphism screening by polymerase chain reaction in Korean patients with rheumatoid arthritis receiving methotrexate. J Rheumatol 2006;33(7):1266–74.

[72] Xiao H, Xu J, Zhou X, Stankovich J, Pan F, Zhang Z, et al. Associations between the genetic polymorphisms of MTHFR and outcomes of methotrexate treatment in rheumatoid arthritis. Clin Exp Rheumatol 2010;28(5):728–33.

[73] van Ede AE, Laan RF, Blom HJ, Huizinga TW, Haagsma CJ, Giesendorf BA, et al. The C677T mutation in the methylenetetrahydrofolate reductase gene: a genetic risk factor for methotrexate-related elevation of liver enzymes in rheumatoid arthritis patients. Arthritis Rheum 2001;44(11):2525–30.

[74] Ranganathan P, Culverhouse R, Marsh S, Mody A, Scott-Horton TJ, Brasington R, et al. Methotrexate (MTX) pathway gene polymorphisms and their effects on MTX toxicity in Caucasian and African American patients with rheumatoid arthritis. J Rheumatol 2008;35(4):572–9.

[75] Caliz R, del Amo J, Balsa A, Blanco F, Silva L, Sanmarti R, et al. The C677T polymorphism in the MTHFR gene is associated with the toxicity of methotrexate in a Spanish rheumatoid arthritis population. Scand J Rheumatol 2012;41(1):10–14.

[76] Horie N, Aiba H, Oguro K, Hojo H, Takeishi K. Functional analysis and DNA polymorphism of the tandemly repeated sequences in the 5'-terminal regulatory region of the human gene for thymidylate synthase. Cell Struct Funct 1995;20(3):191–7.

[77] Kawakami K, Omura K, Kanehira E, Watanabe Y. Polymorphic tandem repeats in the thymidylate synthase gene is associated with its protein expression in human gastrointestinal cancers. Anticancer Res 1999;19(4B):3249–52.

[78] DiPaolo A, Chu E. The role of thymidylate synthase as a molecular biomarker. Clin Cancer Res 2004;10(2):411–2.

[79] Pullarkat ST, Stoehlmacher J, Ghaderi V, Xiong YP, Ingles SA, Sherrod A, et al. Thymidylate synthase gene polymorphism determines response and toxicity of 5-FU chemotherapy. Pharmacogenomics J 2001;1(1):65–70.

[80] Kawakami K, Watanabe G. Identification and functional analysis of single nucleotide polymorphism in the tandem repeat sequence of thymidylate synthase gene. Cancer Res 2003;63(18):6004–7.

[81] Ulrich CM, Bigler J, Velicer CM, Greene EA, Farin FM, Potter JD. Searching expressed sequence tag databases: discovery and confirmation of a common polymorphism in the thymidylate synthase gene. Cancer Epidemiol Biomarkers Prev 2000;9(12):1381–5.

[82] Chan ES, Cronstein BN. Molecular action of methotrexate in inflammatory diseases. Arthritis Res 2002;4(4):266–73.

[83] Sherr CJ. G1 phase progression: cycling on cue. Cell 1994;79(4):551–5.

[84] Betticher DC, Thatcher N, Altermatt HJ, Hoban P, Ryder WD, Heighway J. Alternate splicing produces a novel cyclin D1 transcript. Oncogene 1995;11(5):1005–11.

[85] Lu F, Gladden AB, Diehl JA. An alternatively spliced cyclin D1 isoform, cyclin D1b, is a nuclear oncogene. Cancer Res 2003;63(21):7056–61.

[86] Hochhauser D, Schnieders B, Ercikan-Abali E, Gorlick R, Muise-Helmericks R, Li WW, et al. Effect of cyclin D1 overexpression on drug sensitivity in a human fibrosarcoma cell line. J Natl Cancer Inst 1996;88(18):1269–75.

[87] Davis JP, Cain GA, Pitts WJ, Magolda RL, Copeland RA. The immunosuppressive metabolite of leflunomide is a potent inhibitor of human dihydroorotate dehydrogenase. Biochemistry 1996;35(4):1270–3.

[88] Grabar PB, Rozman B, Logar D, Praprotnik S, Dolzan V. Dihydroorotate dehydrogenase polymorphism influences the toxicity of leflunomide treatment in patients with rheumatoid arthritis. Ann Rheum Dis 2009;68(8):1367–8.

[89] Allikmets R, Gerrard B, Hutchinson A, Dean M. Characterization of the human ABC superfamily: isolation and mapping of 21 new genes using the expressed sequence tags database. Hum Mol Genet 1996;5(10):1649–55.

[90] Cusatis G, Sparreboom A. Pharmacogenomic importance of ABCG2. Pharmacogenomics 2008;9(8):1005–9.

[91] Kim KA, Joo HJ, Park JY. Effect of ABCG2 genotypes on the pharmacokinetics of A771726, an active metabolite of prodrug leflunomide, and association of A771726 exposure with serum uric acid level. Eur J Clin Pharmacol 2011;67(2):129–34.

[92] Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. Arthritis Res Ther 2009;11(1):R7.

[93] Dziedziejko V, Kurzawski M, Safranow K, Chlubek D, Pawlik A. The effect of ESR1 and ESR2 gene polymorphisms on the outcome of rheumatoid arthritis treatment with leflunomide. Pharmacogenomics 2011;12(1):41–7.

[94] Kalgutkar AS, Nguyen HT, Vaz AD, Doan A, Dalvie DK, McLeod DG, et al. *In vitro* metabolism studies on the isoxazole ring scission in the anti-inflammatory agent lefluonomide to its active alpha-cyanoenol metabolite A771726: mechanistic similarities with the cytochrome P450-catalyzed dehydration of aldoximes. Drug Metab Dispos 2003;31(10):1240–50.

[95] Bohanec Grabar P, Rozman B, Tomsic M, Suput D, Logar D, Dolzan V. Genetic polymorphism of CYP1A2 and the toxicity of leflunomide treatment in rheumatoid arthritis patients. Eur J Clin Pharmacol 2008;64(9):871–6.

[96] Bohanec Grabar P, Grabnar I, Rozman B, Logar D, Tomsic M, Suput D, et al. Investigation of the influence of CYP1A2 and CYP2C19 genetic polymorphism on 2-Cyano-3-hydroxy-N-[4-(trifluoromethyl)phenyl]-2-butenamide (A77 1726) pharmacokinetics in leflunomide-treated patients with rheumatoid arthritis. Drug Metab Dispos 2009;37(10):2061–8.

[97] Wiese MD, Schnabl M, O'Doherty C, Spargo LD, Sorich MJ, Cleland LG, et al. Polymorphisms in cytochrome P450 2C19 enzyme and cessation of leflunomide in patients with rheumatoid arthritis. Arthritis Res Ther 2012;14(4):R163.

[98] Stolk JN, Boerbooms AM, de Abreu RA, de Koning DG, van Beusekom HJ, Muller WH, et al. Reduced thiopurine methyltransferase activity and development of side effects of azathioprine treatment in patients with rheumatoid arthritis. Arthritis Rheum 1998;41(10):1858–66.

[99] Marinaki AM, Ansari A, Duley JA, Arenas M, Sumi S, Lewis CM, et al. Adverse drug reactions to azathioprine therapy are associated with polymorphism in the gene encoding inosine triphosphate pyrophosphatase (ITPase). Pharmacogenetics 2004;14(3):181–7.

[100] Krynetski EY, Tai HL, Yates CR, Fessing MY, Loennechen T, Schuetz JD, et al. Genetic polymorphism of thiopurine S-methyltransferase: clinical importance and molecular mechanisms. Pharmacogenetics 1996;6(4):279–90.

[101] Tai HL, Krynetski EY, Yates CR, Loennechen T, Fessing MY, Krynetskaia NF, et al. Thiopurine S-methyltransferase deficiency: two nucleotide transitions define the most prevalent mutant allele associated with loss of catalytic activity in Caucasians. Am J Hum Genet 1996;58(4):694–702.

[102] Tai HL, Krynetski EY, Schuetz EG, Yanishevski Y, Evans WE. Enhanced proteolysis of thiopurine S-methyltransferase (TPMT) encoded by mutant alleles in humans (TPMT*3A, TPMT*2): mechanisms for the genetic polymorphism of TPMT activity. Proc Natl Acad Sci 1997;94(12):6444–9.

[103] Yates CR, Krynetski EY, Loennechen T, Fessing MY, Tai HL, Pui CH, et al. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. Ann Intern Med 1997;126(8):608–14.

[104] Ameyaw MM, Collie-Duguid ES, Powrie RH, Ofori-Adjei D, McLeod HL. Thiopurine methyltransferase alleles in British and Ghanaian populations. Human Mol Genet 1999;8(2):367–70.

[105] Hon YY, Fessing MY, Pui CH, Relling MV, Krynetski EY, Evans WE. Polymorphism of the thiopurine S-methyltransferase gene in African-Americans. Human Mol Genet 1999;8(2):371–6.

[106] Evans WE, Hon YY, Bomgaars L, Coutre S, Holdsworth M, Janco R, et al. Preponderance of thiopurine S-methyltransferase deficiency and heterozygosity among patients intolerant to mercaptopurine or azathioprine. J Clin Oncol 2001;19(8):2293–301.

[107] Black AJ, McLeod HL, Capell HA, Powrie RH, Matowe LK, Pritchard SC, et al. Thiopurine methyltransferase genotype predicts therapy-limiting severe toxicity from azathioprine. Ann Intern Med 1998;129(9):716–8.

[108] Corominas H, Domenech M, Laiz A, Gich I, Geli C, Diaz C, et al. Is thiopurine methyltransferase genetic polymorphism a major factor for withdrawal of azathioprine in rheumatoid arthritis patients? Rheumatology 2003;42(1):40–5.

[109] Boonsrirat U, Angsuthum S, Vannaprasaht S, Kongpunvijit J, Hirankarn N, Tassaneeyakul W, et al. Azathioprine-induced fatal myelosuppression in systemic lupus erythematosus patient carrying TPMT*3C polymorphism. Lupus 2008;17(2):132–4.

[110] Higgs JE, Payne K, Roberts C, Newman WG. Are patients with intermediate TPMT activity at increased risk of myelosuppression when taking thiopurine medications? Pharmacogenomics 2010;11(2):177–88.

[111] von Ahsen N, Armstrong VW, Behrens C, von Tirpitz C, Stallmach A, Herfarth H, et al. Association of inosine triphosphatase 94C>A and thiopurine S-methyltransferase deficiency with adverse events and study drop-outs under azathioprine therapy in a prospective Crohn disease study. Clin Chem 2005;51(12):2282–8.

[112] Shipkova M, Franz J, Abe M, Klett C, Wieland E, Andus T. Association between adverse effects under azathioprine therapy and inosine triphosphate pyrophosphatase activity in patients with chronic inflammatory bowel disease. Ther Drug Monit 2011;33(3):321–8.

[113] Gearry RB, Roberts RL, Barclay ML, Kennedy MA. Lack of association between the ITPA 94C>A polymorphism and adverse effects from azathioprine. Pharmacogenetics 2004;14(11):779–81.

[114] Eklund BI, Moberg M, Bergquist J, Mannervik B. Divergent activities of human glutathione transferases in the bioactivation of azathioprine. Mol Pharmacol 2006;70(2):747–54.

[115] Stocco G, Martelossi S, Barabino A, Decorti G, Bartoli F, Montico M, et al. Glutathione-S-transferase genotypes and the adverse effects of azathioprine in young patients with inflammatory bowel disease. Inflamm Bowel Dis 2007;13(1):57–64.

[116] Stocco G, Cuzzoni E, De Iudicibus S, Franca R, Favretto D, Malusa N, et al. Deletion of glutathione-S-transferase M1 reduces azathioprine metabolite concentrations in young patients with inflammatory bowel disease. J Clin Gastroenterol 2013. Epub ahead of print

[117] Kerstens PJ, Stolk JN, De Abreu RA, Lambooy LH, van de Putte LB, Boerbooms AA. Azathioprine-related bone marrow toxicity and low activities of purine enzymes in patients with rheumatoid arthritis. Arthritis Rheum 1995;38(1):142–5.

[118] Das KM, Eastwood MA, McManus JP, Sircus W. Adverse reactions during salicylazosulfapyridine therapy and the relation with drug metabolism and acetylator phenotype. N Engl J Med 1973;289(10):491–5.

[119] Pullar T, Capell HA. Variables affecting efficacy and toxicity of sulphasalazine in rheumatoid arthritis. A Rev Drugs 1986;32(Suppl 1):54–7.

[120] Tanaka E, Taniguchi A, Urano W, Nakajima H, Matsuda Y, Kitamura Y, et al. Adverse effects of sulfasalazine in patients with rheumatoid arthritis are associated with diplotype configuration at the N-acetyltransferase 2 gene. J Rheumatol 2002;29(12):2492–9.

[121] Wadelius M, Stjernberg E, Wiholm BE, Rane A. Polymorphisms of NAT2 in relation to sulphasalazine-induced agranulocytosis. Pharmacogenetics 2000;10(1):35–41.

[122] Kuhn UD, Anschutz M, Schmucker K, Schug BS, Hippius M, Blume HH. Phenotyping with sulfasalazine - time dependence and relation to NAT2 pharmacogenetics. Int J Clin Pharmacol Ther 2010;48(1):1–10.

[123] Ohkuma S, Poole B. Fluorescence probe measurement of the intralysosomal pH in living cells and the perturbation of pH by various agents. Proc Natl Acad Sci USA 1978;75(7):3327–31.

[124] Weber SM, Levitz SM. Chloroquine interferes with lipopolysaccharide-induced TNF-alpha gene expression by a nonlysosomotropic mechanism. J Immunol 2000;165(3):1534–40.

[125] Wozniacka A, Lesiak A, Narbutt J, McCauliffe DP, Sysa-Jedrzejowska A. Chloroquine treatment influences proinflammatory cytokine levels in systemic lupus erythematosus patients. Lupus 2006;15(5):268–75.

[126] Jang CH, Choi JH, Byun MS, Jue DM. Chloroquine inhibits production of TNF-alpha, IL-1beta and IL-6 from lipopolysaccharide-stimulated human monocytes/macrophages by different modes. Rheumatology (Oxford) 2006;45(6):703–10.

[127] Macfarlane DE, Manzel L. Antagonism of immunostimulatory CpG-oligodeoxynucleotides by quinacrine, chloroquine, and structurally related compounds. J Immunol 1998;160(3):1122–31.

[128] Kim KA, Park JY, Lee JS, Lim S. Cytochrome P450 2C8 and CYP3A4/5 are involved in chloroquine metabolism in human liver microsomes. Arch Pharm Res 2003;26(8):631–7.

[129] Projean D, Baune B, Farinotti R, Flinois JP, Beaune P, Taburet AM, et al. *In vitro* metabolism of chloroquine: identification of CYP2C8,

CYP3A4, and CYP2D6 as the main isoforms catalyzing N-desethylchloroquine formation. Drug Metab Dispos 2003;31(6):748–54.

[130] Parikh S, Ouedraogo JB, Goldstein JA, Rosenthal PJ, Kroetz DL. Amodiaquine metabolism is impaired by common polymorphisms in CYP2C8: implications for malaria treatment in Africa. Clin Pharmacol Ther 2007;82(2):197–203.

[131] Wahie S, Daly AK, Cordell HJ, Goodfield MJ, Jones SK, Lovell CR, et al. Clinical and pharmacogenetic influences on response to hydroxychloroquine in discoid lupus erythematosus: a retrospective cohort study. J Invest Dermatol 2011;131(10):1981–6.

[132] Suarez A, Castro P, Alonso R, Mozo L, Gutierrez C. Interindividual variations in constitutive interleukin-10 messenger RNA and protein levels and their association with genetic polymorphisms. Transplantation 2003;75(5):711–7.

[133] Suarez A, Lopez P, Mozo L, Gutierrez C. Differential effect of IL10 and TNF {alpha} genotypes on determining susceptibility to discoid and systemic lupus erythematosus. Ann Rheum Dis 2005;64(11):1605–10.

[134] Turner DM, Williams DM, Sankaran D, Lazarus M, Sinnott PJ, Hutchinson IV. An investigation of polymorphism in the interleukin-10 gene promoter. Eur J Immunogenet 1997;24(1):1–8.

[135] Seitz M, Wirthmuller U, Moller B, Villiger PM. The -308 tumour necrosis factor-alpha gene polymorphism predicts therapeutic response to TNFalpha-blockers in rheumatoid arthritis and spondyloarthritis patients. Rheumatology (Oxford) 2007;46(1):93–6.

[136] Cuchacovich M, Ferreira L, Aliste M, Soto L, Cuenca J, Cruzat A, et al. Tumour necrosis factor-alpha (TNF-alpha) levels and influence of -308 TNF-alpha promoter polymorphism on the responsiveness to infliximab in patients with rheumatoid arthritis. Scand J Rheumatol 2004;33(4):228–32.

[137] Lopez P, Gomez J, Mozo L, Gutierrez C, Suarez A. Cytokine polymorphisms influence treatment outcomes in SLE patients treated with antimalarial drugs. Arthritis Res Ther 2006;8(2):R42.

[138] Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, et al. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. N Engl J Med 2000;343(22):1586–93.

[139] Genovese MC, Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, et al. Etanercept versus methotrexate in patients with early rheumatoid arthritis: two-year radiographic and clinical outcomes. Arthritis Rheum 2002;46(6):1443–50.

[140] Maini RN, Breedveld FC, Kalden JR, Smolen JS, Davis D, Macfarlane JD, et al. Therapeutic efficacy of multiple intravenous infusions of anti-tumor necrosis factor alpha monoclonal antibody combined with low-dose weekly methotrexate in rheumatoid arthritis. Arthritis Rheum 1998;41(9):1552–63.

[141] Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS, et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. Arthritis Rheum 2004;50(5):1400–11.

[142] Pavy S, Toonen EJ, Miceli-Richard C, Barrera P, van Riel PL, Criswell LA, et al. Tumour necrosis factor alpha -308G->A polymorphism is not associated with response to TNFalpha blockers in Caucasian patients with rheumatoid arthritis: systematic review and meta-analysis. Ann Rheum Dis 2010;69(6):1022–8.

[143] Lee YH, Ji JD, Bae SC, Song GG. Associations between tumor necrosis factor-alpha (TNF-alpha) -308 and -238 G/A polymorphisms and shared epitope status and responsiveness to TNF-alpha blockers in rheumatoid arthritis: a metaanalysis update. J Rheumatol 2010;37(4):740–6.

[144] Zeng Z, Duan Z, Zhang T, Wang S, Li G, Gao J, et al. Association between tumor necrosis factor-alpha (TNF-alpha) promoter -308 G/A and response to TNF-alpha blockers in rheumatoid arthritis: a meta-analysis. Mod Rheumatol 2013;23(3):489–95.

[145] Maxwell JR, Potter C, Hyrich KL, Barton A, Worthington J, Isaacs JD, et al. Association of the tumour necrosis factor-308 variant with differential response to anti-TNF agents in the treatment of rheumatoid arthritis. Hum Mol Genet 2008;17(22):3532–8.

[146] Fabris M, Tolusso B, Di Poi E, Assaloni R, Sinigaglia L, Ferraccioli G. Tumor necrosis factor-alpha receptor II polymorphism in patients from southern Europe with mild-moderate and severe rheumatoid arthritis. J Rheumatol 2002;29(9):1847–50.

[147] Criswell LA, Lum RF, Turner KN, Woehl B, Zhu Y, Wang J, et al. The influence of genetic variation in the HLA-DRB1 and LTA-TNF regions on the response to treatment of early rheumatoid arthritis with methotrexate or etanercept. Arthritis Rheum 2004;50(9):2750–6.

[148] Ongaro A, De Mattei M, Pellati A, Caruso A, Ferretti S, Masieri FF, et al. Can tumor necrosis factor receptor II gene 676T>G polymorphism predict the response grading to anti-TNFalpha therapy in rheumatoid arthritis? Rheumatol Int 2008;28(9):901–8.

[149] Martinez A, Salido M, Bonilla G, Pascual-Salcedo D, Fernandez-Arquero M, de Miguel S, et al. Association of the major histocompatibility complex with response to infliximab therapy in rheumatoid arthritis patients. Arthritis Rheum 2004;50(4):1077–82.

[150] Tutuncu Z, Kavanaugh A, Zvaifler N, Corr M, Deutsch R, Boyle D. Fcgamma receptor type IIIA polymorphisms influence treatment outcomes in patients with inflammatory arthritis treated with tumor necrosis factor alpha-blocking agents. Arthritis Rheum 2005;52(9):2693–6.

[151] Canete JD, Suarez B, Hernandez MV, Sanmarti R, Rego I, Celis R, et al. Influence of variants of Fc gamma receptors IIA and IIIA on the American College of Rheumatology and European League Against Rheumatism responses to anti-tumour necrosis factor alpha therapy in rheumatoid arthritis. Ann Rheum Dis 2009;68(10):1547–52.

[152] Tsukahara S, Ikari K, Sato E, Yamanaka H, Hara M, Tomatsu T, et al. A polymorphism in the gene encoding the Fcgamma IIIA receptor is a possible genetic marker to predict the primary response to infliximab in Japanese patients with rheumatoid arthritis. Ann Rheum Dis 2008;67(12):1791–2.

[153] Cui J, Saevarsdottir S, Thomson B, Padyukov L, van der Helm-van Mil AH, Nititham J, et al. Rheumatoid arthritis risk allele PTPRC is also associated with response to anti-tumor necrosis factor alpha therapy. Arthritis Rheum 2010;62(7):1849–61.

[154] Plant D, Prajapati R, Hyrich KL, Morgan AW, Wilson AG, Isaacs JD, et al. Replication of association of the PTPRC gene with response to anti-tumor necrosis factor therapy in a large UK cohort. Arthritis Rheum 2012;64(3):665–70.

[155] Bowes JD, Potter C, Gibbons LJ, Hyrich K, Plant D, Morgan AW, et al. Investigation of genetic variants within candidate genes of the TNFRSF1B signalling pathway on the response to anti-TNF agents in a UK cohort of rheumatoid arthritis patients. Pharmacogenet Genomics 2009;19(4):319–23.

[156] Coulthard LR, Taylor JC, Eyre S, Robinson JI, Wilson AG, Isaacs JD, et al. Genetic variants within the MAP kinase signalling network and anti-TNF treatment response in rheumatoid arthritis patients. Ann Rheum Dis 2011;70(1):98–103.

[157] Tan RJ, Gibbons LJ, Potter C, Hyrich KL, Morgan AW, Wilson AG, et al. Investigation of rheumatoid arthritis susceptibility genes identifies association of AFF3 and CD226 variants with response to anti-tumour necrosis factor treatment. Ann Rheum Dis 2010;69(6):1029–35.

[158] Schotte H, Schluter B, Drynda S, Willeke P, Tidow N, Assmann G, et al. Interleukin 10 promoter microsatellite polymorphisms are associated with response to long term treatment with etanercept in patients with rheumatoid arthritis. Ann Rheum Dis 2005;64(4):575–81.

[159] Padyukov L, Lampa J, Heimburger M, Ernestam S, Cederholm T, Lundkvist I, et al. Genetic markers for the efficacy of tumour necrosis factor blocking therapy in rheumatoid arthritis. Ann Rheum Dis 2003;62(6):526–9.

[160] Jancic I, Arsenovic-Ranin N, Sefik-Bukilica M, Zivojinovic S, Damjanov N, Spasovski V, et al. −174G/C interleukin-6 gene promoter polymorphism predicts therapeutic response to etanercept in rheumatoid arthritis. Rheumatol Int 2013;33(6):1481–6.

[161] Santee SM, Owen-Schaub LB. Human tumor necrosis factor receptor p75/80 (CD120b) gene structure and promoter characterization. J Biol Chem 1996;271(35):21151–9.

[162] Morita C, Horiuchi T, Tsukamoto H, Hatta N, Kikuchi Y, Arinobu Y, et al. Association of tumor necrosis factor receptor type II polymorphism 196R with Systemic lupus erythematosus in the Japanese: molecular and functional analysis. Arthritis Rheum 2001;44(12):2819–27.

[163] Vasilopoulos Y, Bagiatis V, Stamatopoulou D, Zisopoulos D, Alexiou I, Sarafidou T, et al. Association of anti-CCP positivity and carriage of TNFRII susceptibility variant with anti-TNF-alpha response in rheumatoid arthritis. Clin Exp Rheumatol 2011;29(4):701–4.

[164] Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;30(11):1205–13.

[165] Waldron-Lynch F, Adams C, Amos C, Zhu DK, McDermott MF, Shanahan F, et al. Tumour necrosis factor 5' promoter single nucleotide polymorphisms influence susceptibility to rheumatoid arthritis (RA) in immunogenetically defined multiplex RA families. Genes Immun 2001;2(2):82–7.

[166] Mulcahy B, Waldron-Lynch F, McDermott MF, Adams C, Amos CI, Zhu DK, et al. Genetic variability in the tumor necrosis factor-lymphotoxin region influences susceptibility to rheumatoid arthritis. Am J Hum Genet 1996;59(3):676–83.

[167] Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 2010;42(6):508–14.

[168] Krintel SB, Palermo G, Johansen JS, Germer S, Essioux L, Benayed R, et al. Investigation of single nucleotide polymorphisms and biological pathways associated with response to TNF alpha inhibitors in patients with rheumatoid arthritis. Pharmacogenet Genomics 2012;22(8):577–89.

[169] Barton A, Eyre S, Ke X, Hinks A, Bowes J, Flynn E, et al. Identification of AF4/FMR2 family, member 3 (AFF3) as a novel rheumatoid arthritis susceptibility locus and confirmation of two further pan-autoimmune susceptibility genes. Hum Mol Genet 2009;18(13):2518–22.

[170] Hafler JP, Maier LM, Cooper JD, Plagnol V, Hinks A, Simmonds MJ, et al. CD226 Gly307Ser association with multiple autoimmune diseases. Genes Immun 2009;10(1):5–10.

[171] Marotte H, Pallot-Prades B, Grange L, Tebib J, Gaudin P, Alexandre C, et al. The shared epitope is a marker of severity associated with selection for, but not with response to, infliximab in a large rheumatoid arthritis population. Ann Rheum Dis 2006;65(3):342–7.

[172] Potter C, Cordell HJ, Barton A, Daly AK, Hyrich KL, Mann DA, et al. Association between anti-tumour necrosis factor treatment response and genetic variants within the TLR and NF{kappa}B signalling pathways. Ann Rheum Dis 2010;69(7):1315–20.

[173] Kastbom A, Bratt J, Ernestam S, Lampa J, Padyukov L, Soderkvist P, et al. Fcgamma receptor type IIIA genotype and response to tumor necrosis factor alpha-blocking agents in patients with rheumatoid arthritis. Arthritis Rheum 2007;56(2):448–52.

[174] Liu C, Batliwalla F, Li W, Lee A, Roubenoff R, Beckman E, et al. Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis. Mol Med 2008;14(9–10):575–81.

[175] Plant D, Bowes J, Potter C, Hyrich KL, Morgan AW, Wilson AG, et al. Genome-wide association study of genetic predictors of anti-tumor necrosis factor treatment efficacy in rheumatoid arthritis identifies associations with polymorphisms at seven loci. Arthritis Rheum 2011;63(3):645–53.

[176] Umicevic Mirkov M, Cui J, Vermeulen SH, Stahl EA, Toonen EJ, Makkinje RR, et al. Genome-wide association analysis of anti-TNF drug response in patients with rheumatoid arthritis. Ann Rheum Dis 2013;72(8):1375–81.

[177] Martin M, Romero X, de la Fuente MA, Tovar V, Zapater N, Esplugues E, et al. CD84 functions as a homophilic adhesion molecule and enhances IFN-gamma secretion: adhesion is mediated by Ig-like domain 1. J Immunol 2001;167(7):3668–76.

[178] Tangye SG, Nichols KE, Hare NJ, van de Weerdt BC. Functional requirements for interactions between CD84 and Src homology 2 domain-containing proteins and their contribution to human T cell activation. J Immunol 2003;171(5):2485–95.

[179] Cui J, Stahl EA, Saevarsdottir S, Miceli C, Diogo D, Trynka G, et al. Genome-wide association study and gene expression analysis identifies CD84 as a predictor of response to etanercept therapy in rheumatoid arthritis. PLoS Genet 2013;9(3):e1003394.

[180] Acosta-Colman I, Palau N, Tornero J, Fernandez-Nebro A, Blanco F, Gonzalez-Alvaro I, et al. GWAS replication study confirms the association of PDE3A-SLCO1C1 with anti-TNF therapy response in rheumatoid arthritis. Pharmacogenomics 2013;14(7):727–34.

[181] Daien CI, Fabre S, Rittore C, Soler S, Daien V, Tejedor G, et al. TGF beta1 polymorphisms are candidate predictors of the clinical response to rituximab in rheumatoid arthritis. Joint Bone Spine 2012;79(5):471–5.

[182] Coenen MJ, Gregersen PK. Rheumatoid arthritis: a view of the current genetic landscape. Genes Immun 2009;10(2):101–11.

[183] Boissier MC. Cell and cytokine imbalances in rheumatoid synovitis. Joint Bone Spine 2011;78(3):230–4.

[184] Guo Z, Binswanger U, Knoflach A. Role of codon 10 and codon 25 polymorphisms on TGF-beta 1 gene expression and protein synthesis in stable renal allograft recipients. Transplant Proc 2002;34(7):2904–6.

[185] Ruyssen-Witrand A, Rouanet S, Combe B, Dougados M, Le Loet X, Sibilia J, et al. Association between -871C>T promoter polymorphism in the B-cell activating factor gene and the response to rituximab in rheumatoid arthritis patients. Rheumatology (Oxford) 2013;52(4):636–41.

[186] Fabris M, Quartuccio L, Vital E, Pontarini E, Salvin S, Fabro C, et al. The TTTT B lymphocyte stimulator promoter haplotype is

associated with good response to rituximab therapy in seropositive rheumatoid arthritis resistant to tumor necrosis factor blockers. Arthritis Rheum 2013;65(1):88–97.

[187] Ruyssen-Witrand A, Rouanet S, Combe B, Dougados M, Le Loet X, Sibilia J, et al. Fcgamma receptor type IIIA polymorphism influences treatment outcomes in patients with rheumatoid arthritis treated with rituximab. Ann Rheum Dis 2012;71(6):875–7.

[188] Quartuccio L, Fabris M, Pontarini E, Salvin S, Zabotti A, Benucci M, et al. The 158VV Fcgamma receptor 3A genotype is associated with response to rituximab in rheumatoid arthritis: results of an Italian multicentre study. Ann Rheum Dis 2013. [Epub ahead of print].

[189] Kastbom A, Coster L, Arlestig L, Chatzidionysiou A, van Vollenhoven RF, Padyukov L, et al. Influence of FCGR3A genotype on the therapeutic response to rituximab in rheumatoid arthritis: an observational cohort study. BMJ Open 2012;2(5).

[190] Sarsour K, Greenberg J, Johnston JA, Nelson DR, O'Brien LA, Oddoux C, et al. The role of the FcGRIIIa polymorphism in modifying the association between treatment and outcome in patients with rheumatoid arthritis treated with rituximab versus TNF-alpha antagonist therapies. Clin Exp Rheumatol 2013;31(2):189–94.

[191] Fabris M, Quartuccio L, Lombardi S, Saracco M, Atzeni F, Carletto A, et al. The CC homozygosis of the -174G>C IL-6 polymorphism predicts a lower efficacy of rituximab therapy in rheumatoid arthritis. Autoimmun Rev 2012;11(5):315–20.

[192] Wang J, Bansal AT, Martin M, Germer S, Benayed R, Essioux L, et al. Genome-wide association analysis implicates the involvement of eight loci with response to tocilizumab for the treatment of rheumatoid arthritis. Pharmacogenomics J 2013;13(3):235–41.

[193] Ranganathan P, Culverhouse R, Marsh S, Ahluwalia R, Shannon WD, Eisen S, et al. Single nucleotide polymorphism profiling across the methotrexate pathway in normal subjects and patients with rheumatoid arthritis. Pharmacogenomics 2004;5(5):559–69.

[194] Martinez A, Salido M, Bonilla G, Pascual-Salcedo D, Fernandez-Arquero M, de Miguel S, et al. Association of the major histocompatibility complex with response to infliximab therapy in rheumatoid arthritis patients. Arthritis Rheum 2004;50(4):1077–82.

[195] Veenstra DL, Higashi MK, Phillips KA. Assessing the cost-effectiveness of pharmacogenomics. Aaps Pharmsci 2000;2(3):E29.

This page is intentionally left blank

Chapter 30

Pharmacogenomics of Oral Antidiabetic Drugs

Vangelis G. Manolopoulos and Georgia Ragia
Laboratory of Pharmacology and Clinical Pharmacology, Medical School, Democritus University of Thrace, Alexandroupolis, Greece

Chapter Outline

30.1 Overview 683
30.2 Genetic Basis of Diabetes 683
30.3 Management of Diabetes 684
30.4 The Need for Pharmacogenomics and Stratified Medicine 685
30.5 Genetic Variants Associated with Response to Oral Hypoglycemic Agents 685
  30.5.1 Biguanides (Metformin) 685
  30.5.2 Sulfonylureas 703
  30.5.3 Nonsulfonylurea Insulin Secretagogues (Glinides) 706
  30.5.4 Thiazolidinediones 707
30.6 Stratified Oral Antidiabetic Medicine: Pharmacogenomics Application 709
30.7 Conclusions: Future Perspectives 710
Abbreviations 710
References 710

30.1 OVERVIEW

Diabetes is one of the most common chronic diseases and has been steadily gaining in prevalence worldwide in recent decades. According to the fifth edition of the *Diabetes Atlas* (2012 update) 371 million people have diabetes and 187 million are still to be diagnosed. This upward trend will continue (www.idf.org/diabetesatlas). Approximately 90−95% of all diabetic individuals present with type 2 diabetes mellitus (T2DM).

T2DM is characterized by hyperglycemia, relative impairment in insulin secretion by β-pancreatic cells, and decreased tissue responsiveness to insulin (insulin resistance). Most common T2DM complications due to long-term hyperglycemia are micro- and macrovascular and neuropathic complications, such as retinopathy, nephropathy, coronary artery disease, peripheral vascular disease, stroke, and neuropathy [1]. Because of these complications, T2DM has a tremendous impact on healthcare systems and constitutes a major cause of mortality in the Western world. It has been estimated that, by the end of 2012, 4.8 million people had died from diabetes-related complications and that half of these deaths were people under the age of 60.

T2DM is one of the most complicated diseases, implicating several pathophysiologic mechanisms: oxidative stress, endoplasmic reticulum stress (ER stress) in B cells, and tissue inflammation; more recently, autoimmune mediators that cause inflammation have been associated with decline in insulin secretion [2]. Furthermore, T2DM incidence is attributed to both demographic and lifestyle factors, such as obesity, aging, physical inactivity, and genetic factors.

30.2 GENETIC BASIS OF DIABETES

The genomics of T2DM is an evolving field. Accumulated data on T2DM genomics research highlight its polygenic origin. Several gene polymorphisms have been associated in case-control association studies with T2DM incidence [3]. Case-control association studies are in principle conducted on the basis of hypotheses about the involvement of specific biologic factors with T2DM pathogenesis; these may range from factors directly regulating B cell function [4] to components of other systems such as the serotonergic system [5].

In addition, more than 35 genome-wide association studies (GWASs) on T2DM have been published as of

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00030-X
© 2014 Elsevier Inc. All rights reserved.

mid-2013 and can be found at www.genome.gov/gwastudies/index.cfm?pageid=26525384#searchForm. These hypothesis-free tools have revealed the association of several gene polymorphisms with T2DM incidence. It should be noted that results from GWASs have identified polymorphisms with as yet unknown functions relating to gene regulation or protein function [6]. In any case, the common theme that has emerged from GWASs confirms what was already known about the polygenic origin of complex diseases—that is, multiple gene polymorphisms confer only a modest increase in T2DM risk.

The genetic imprint of T2DM is being gradually discovered as the association of several gene polymorphisms with T2DM incidence is being replicated in both GWASs and case-control studies. A description of these studies is out of the scope of this chapter; interested readers are referred to a recent comprehensive review on the subject [3]. The most replicated gene associated with T2DM encodes transcription factor 7-like 2 (*TCF7L2*) that is expressed in the mature and developing β-pancreatic cells. This transcription factor is involved in the canonical β-catenin-dependent Wnt signaling pathway, which in turn regulates cell differentiation, proliferation, migration and survival, and organ development. *TCF7L2* is commonly called the diabetes gene because of its consistent association with T2DM development. It plays an important role in several vital functions in the pancreatic islet, such as pancreas development, B cell mass, insulin processing, and the maintenance of secretory function in mature B cells [7].

The *TCF7L2* gene single-nucleotide polymorphism (SNP) rs7903146 (*C*>*T*) strongly predicts the incidence of T2DM in populations of different ancestry (European, Asian, and African), and it is nearly in complete linkage disequilibrium (LD) with other variants of *TCF7L2*, such as rs12255372 (*G*>*T*), that are also associated with T2DM. The association of each SNP with T2DM incidence correlates 92% of the time with the other. The *TCF7L2* rs7903146 (*C*>*T*) SNP is among the most highly replicated and consistent associations with a disease of multigenic origin; its association with T2DM derives from consistent results of GWASs and global meta-analyses on different quantitative glycemic traits such as fasting plasma glucose levels and glycated hemoglobin (HbA1c). The attributable risk in disease incidence is driven by the *TCF7L2* rs7903146 *T* allele and confers an odds ratio (OR) of 1.46 for developing T2DM—a value similar in all human populations tested for this variant allele [8].

Though *TCF7L2* harbors the variant with the strongest effect on T2DM, the molecular mechanism of variation in this gene which increases the risk of T2DM remains elusive. The genetic variation of *TCF7L2*, rs7903146, is located in a noncoding region and is assumed to alter gene expression. More recent data indicate that *TCF7L2* rs7903146 is located at an islet-selective open chromatin site and consequently alters enhancer activity [7]. Its predominant direct or indirect role in B cell function and possible stimulation of B cell proliferation have also been suggested. (Other gene polymorphisms associated with T2DM in GWASs can be found in detail at the website of the National Human Genome Research Institute, www.genome.gov/gwastudies/index.cfm?pageid=26525384#searchForm.). It is widely accepted that numerous genetic polymorphisms contribute to T2DM causality, and it is believed that combined analysis of polymorphic gene locus confers increased risk [3]. Although gene polymorphisms are crucial for understanding human biology, this static aspect of genomic information is limited for explaining interindividual differences in gene expression, disease susceptibility, and drug response.

In this era, in which the entire genome sequence has been decoded, there is still a gap in knowledge about the hereditability of multigenic diseases such as T2DM, which has led to broader research on the identification of T2DM predisposing factors beyond the genome. The nonplasticity of the genome in terms of nucleotide sequence has led to the molecular characterization of other nongenetic alterations that affect gene expression [9]. The dynamic complexity of the human genome, as well as that of the patterns of altered gene expression in target tissues due to endogenous modulators or environmental factors, has opened the way to a new field of research, *epigenomics*.

Epigenomics focuses on changes in gene function that derive not from changes in the DNA sequence, but from reversible regulation of transcriptional control that dictates how genes are expressed. The general concept of epigenomics lies in the molecular basis of DNA organization. Epigenetic patterns include modifications of DNA wrapping that rule out the availability of a gene for transcription, and therefore expression, and mainly consist of three distinct molecular mechanisms of gene transcription and translation regulation: methylation; acetylation; and microRNAs.

During the last ten years, intense research efforts have been devoted to the study of epigenetic modifications, the factors that promote them, and the challenges remaining in implementing them as prognostic markers or as novel therapeutic targets in T2DM. It should be noted that most suggested mechanisms and implications of epigenomic involvement in T2DM development and diabetic complications are still at the hypothesis level. However, studies of human tissues or cells are beginning to emerge and will provide a more solid basis for understanding the mechanisms involved. Discussion of these studies is beyond the scope of this chapter, and interested readers are directed to a recent comprehensive review on the subject [10].

30.3 MANAGEMENT OF DIABETES

Management of T2DM consists initially of lifestyle modifications (diet and exercise) and eventually of pharmacologic treatment, which sooner or later becomes essential in most T2DM patients. T2DM pharmacologic treatment aims at achieving specific glycemic goals to substantially reduce the risk of both vasculature complications and morbidity. Measurement of HbA1c levels primarily identifies the average plasma glucose concentration over prolonged periods of time and serves as a marker for average blood glucose levels over the months prior to measurement. Guidelines for T2DM management released by the American Diabetes Association (ADA) in 2009 required that HbA1c levels have to be maintained below 7% [11]. More recently, ADA and the European Association for the Study of Diabetes (EASD) updated the recommendations for antihyperglycemic therapy in T2DM adults, stressing the fact that the best available evidence on glycemic control should be incorporated in a patient-centered, multifactorial risk reduction framework [12].

The main classes of oral drugs that are currently available to treat T2DM include biguanide metformin, sulfonylureas, nonsulfonylurea insulin secretagogues (glinides), thiazolidinediones, α-glucosidase inhibitors, and dipeptidyl peptidase IV inhibitors. Newer drug classes such as bile acid resins and dopamine agonists have also gained U.S. Food and Drug Administration (FDA) approval for the treatment of diabetes. Combination therapies are a common approach in T2DM patients, since during disease progression, most patients may require more than one antidiabetic drug to achieve glycemic goals and restrict hyperglycemia [11]. However, despite major recent advances in diabetes care, a large proportion of patients remain in poor glycemic control.

Throughout the years, a consistent observation in antidiabetic drug treatment has been that significant variability exists in drug disposition, glycemic response, tolerability, and incidence of adverse effects, even in T2DM patients who have apparent requirements for antidiabetic regimens. Among other factors, interindividual variability in patient responses to antidiabetic drugs is affected by gene polymorphisms. Pharmacogenomics studies individual pharmacologic response based on genotype [13]. Clinical application of gene polymorphisms in components that affect the pharmacokinetics of antidiabetic drugs, that are involved in their therapeutic mechanisms, and/or that cause adverse effects has the potential to improve T2DM management by helping clinicians effectively prescribe these medications [10].

Pharmacogenomic analysis is an additional tool that can be used, together with HbA1c monitoring—the current gold standard for glycemic control—to improve the choice and dosing of oral antidiabetic treatment by individualizing drug therapy; it does not challenge HbA1c monitoring but rather functions in a complementary fashion because it is addressing a step prior to monitoring blood glucose levels.

30.4 THE NEED FOR PHARMACOGENOMICS AND STRATIFIED MEDICINE

The study of genetic variation in association with response to oral antidiabetic medications focuses on genes involved in drug pharmacokinetics and pharmacodynamics, as well as on genes that are involved in the underlying disease condition or insulin secretion pathway. Most available pharmacogenomic data on antidiabetic medications derive from studies on early drugs, including biguanide (metformin), sulfonylureas, nonsulfonylurea insulin secretagogues, and thiazolidinediones. To the best of our knowledge, pharmacogenetic studies are not yet available on the remaining drug classes used for T2DM treatment.

Information on all gene polymorphisms that have been studied in association with drug response—their chromosome location, effect on encoded protein, and the respective antidiabetic medication they affect—is presented in Table 30.1. To further help the reader in easily accessing the published studies, all currently available data on gene polymorphisms associated with antidiabetic drug response are briefly listed in three tables: Table 30.2 presents polymorphisms in genes mediating the metabolism of antidiabetic drugs; Table 30.3 (later in the chapter) presents polymorphisms in genes encoding drug receptors and drug transporters; Table 30.4 presents polymorphisms in genes of all other molecules involved in disease pathogenesis. Furthermore, given the significance of clinical trials assessing implementation of pharmacogenomics in routine clinical practice, completed and ongoing clinical trials on antidiabetic drug pharmacogenomics are listed in Table 30.5 (also see www.clinicaltrials.gov). All tables can be found in the Appendix.

30.5 GENETIC VARIANTS ASSOCIATED WITH RESPONSE TO ORAL HYPOGLYCEMIC AGENTS

30.5.1 Biguanides (Metformin)

Metformin is the only available biguanide currently in clinical use; the other one, phenformin, was discontinued because of adverse events. Metformin is well tolerated by the majority of patients and acts by inhibiting excessive hepatic glucose production (gluconeogenesis), and subsequently reduces intestinal glucose absorption and fasting glycemia, and improves insulin sensitivity and peripheral glucose uptake and utilization [14]. In T2DM management, metformin is recommended as a first-line treatment, except when contraindicated. Overall, metformin monotherapy lowers HbA1c by 1.5–3%; nonetheless, many patients require additional glucose control, and metformin is approved for use in combination with other oral antidiabetic agents such as sulfonylureas and thiazolidinediones.

TABLE 30.1 Summary of Genetic Loci and Gene Polymorphisms

Gene (Protein)	Genomic Locus	SNP[a]	SNP Gene Location	Amino Acid Substitution	Effect on Encoded Product	MAF	Association
SLC22A1 (OCT1)	6q25.3	rs12208357*C>T* (c.181*C>T*)	Exon 1	R61C	Reduced activity	0.026	Metformin
		rs34130495*G>A* (c.1201G>A)	Exon 7	G401S	Reduced activity	0.009	Metformin
		rs72552763*GAT del* (c.1260_1262*del GAT*)	Exon 7	M420del	Reduced activity		Metformin
		rs34059508*G>A>C* (c.1393*G>A>C*)	Exon 9	G465R	Reduced activity	0.008	Metformin
		rs622342*A>C* (c.1386-2964*C>A*)	Intron 8	—	Reduced activity	0.271	Metformin
		rs2282143*C>T* (c.1022*C>T*)	Exon 6	P341L	Reduced activity	0.057	Metformin
SLC22A2 (OCT2)	6q25.3	MPJ6_OC2003 (596*C>T*)	Exon 3	T199I	Altered activity	—	Metformin
		MPJ6_OC2004 (602*C>T*)	Exon 3	T201M	Altered activity	—	Metformin
		rs316019*G>T* (808*G>T*)	Exon 4	A270S	Altered activity	0.122	Metformin
SLC47A1 (MATE1)	17p11.2	rs2289669*G>A* (c.922-158*G>A*)	Intron 10	—	Reduced activity	0.343	Metformin
		rs8065082*C>T* (c.1106+1385*C>T*)	Intron 12	—	Reduced activity	0.444	Metformin
ATM (ATM)	11q22-q23	rs11212617*C>A* (c175-5285*G>T*)	Intronic variant	—	Unknown	0.471	Metformin
CYP2C9 (CYP2C9)	10q24	rs179853*C>T* (c.430*C>T*) (**2* allele)	Exon 3	R144C	Reduced activity	0.068	Sulfonylureas, nateglinide
		rs1057910*A>C* (c.1075*A>C*) (**3* allele)	Exon 7	I359L	Reduced activity	0.042	Sulfonylureas, nateglinide
ABCC8 (SUR1)	11q15.1	rs757110 (c.4105*G>T*)	Exon 33	S1369A	Altered activity	0.286	Sulfonylureas
KCNJ11 (Kir6.2)	11p15.1	rs5219 (c.67*A>G*)	Exon 1	E23K	Altered activity	0.274	Sulfonylureas, repaglinide
TCF7L2 (TCF7L2)	10q25.3	rs7903146*C>T* (c.382-41435*C>T*)	Intronic variant	—	Decreased B cell function	0.218	Sulfonylureas
CYP2C8 (CYP2C8)	10q23.33	rs11572103*A>T* (c.805*A>T*) (**2* allele)	Exon 5	I269F	Decreased activity	0.039	Pioglitazone

TABLE 30.1 Continued

Gene (Protein)	Genomic Locus	SNP[a]	SNP Gene Location	Amino Acid Substitution	Effect on Encoded Product	MAF	Association
		rs11572080A>G (c.416G>A) (*3 allele)	Exon 3	R139K	Decreased activity	0.065	Repaglinide, pioglitazone, rosiglitazone
		rs78637571G>T (c.820G>T) (*11 allele)	Exon 6	E274Ter	Decreased activity	0.001	Rosiglitazone
SLCO1B1 (OATP1B1)	12p12.2	rs4149056T>C (c.521T>C)	Exon 5	V174A	Decreased activity	0.123	Repaglinide, nateglinide, pioglitazone, rosiglitazone
		rs2306283A>G (c.388A>G)	Exon 5	N130D	Increased activity	0.405	Repaglinide, nateglinide, pioglitazone
PPARγ (PPARγ)	3p25.2	rs1801282 (c.34C>G)	Exon B	P12A	Transcription reduction	0.066	Pioglitazone, rosiglitazone, troglitazone

[a] *As reported in the dbSNP Short Genetic Variations database (www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs).*

Note: These loci/polymorphisms have been examined in association with oral antidiabetic drug response. The table summarizes their chromosome location, effect on encoded protein, MAF (minor allele frequency), and the respective antidiabetic drug they affect.

Metformin is not metabolized by CYP450 hepatic enzymes, nor does it bind to plasma proteins, but it is excreted unchanged in urine [14]. The drug is widely distributed in body tissues, including intestine, liver, and kidney, by at least two organic cation transporters (OCTs), OCT1 and OCT2, and it is also a substrate of human multidrug and toxin extrusion 1 (MATE1) [15]. OCT1 mediates hepatic uptake of metformin from the circulation, whereas OCT2 and MATE1 facilitate its renal secretion. There is much interindividual variability in metformin pharmacokinetics as measured by differences in trough steady-state metformin plasma concentration. Polymorphisms in both OCT1 and OCT2, encoded by the *SLC22A1* and *SLC22A2* genes, respectively, and in MATE1, encoded by the *SLCO47A1* gene, have been consistently associated with metformin pharmacokinetics and therapeutic response, whereas evidence exists that polymorphisms in other genes potentially modify response to therapy (Tables 30.3 and 30.4).

30.5.1.1 SLC22A1 Gene Polymorphisms and Metformin Response

The *SLC22A1* gene encodes for OCT1 and is located on chromosome 6q25.3; it consists of 11 exons spanning 37 kb [16]. Human *SLC22A1* is highly polymorphic and its numerous polymorphisms have been described in various populations (www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=6580). The most common polymorphisms, their minor allele frequency (MAF), and their effect on transporter activity are shown in Table 30.1.

Research on the association of OCT gene polymorphisms with metformin response began in 2007, when Shu et al. showed that, in $Oct1^{-/-}$ mice, the glucose-lowering effect of metformin was abolished [17]. In the same study, the effect of four variants of OCT1 (*R61C*, *G401S*, *M420del*, and *G465R*), which are associated with reduced metformin transport, was assessed in regard to therapeutic response to metformin in 20 healthy volunteers given two subsequent metformin doses (1000 mg in the evening and 850 mg in the morning). Defective variants were associated with significantly decreased glucose-lowering response to metformin [17] and, in a following study in the same cohort population, with significantly higher area under the concentration–time curve (AUC) and maximum plasma concentration after two oral doses [18].

In a subsequent study of 103 healthy male volunteers, the same OCT1 variants (*R61C*, *G401S*, *M420del*, and *G465R*) were associated with increased metformin renal clearance and decreased hepatic uptake, suggesting that renal OCT1 expression may be an important determinant of metformin response [19].

TABLE 30.2 Gene Polymorphisms of Cytochrome P450 Enzymes

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
Metformin	Metformin not metabolized by CYP450 hepatic enzymes	—	—	—	—	—
Sulfonylureas	*CYP2C9*	*CYP2C9*2* and **3* compared with *CYP2C9*1/*1*	16 healthy volunteers	Plasma concentration of glyburide and glimepiride	*CYP2C9*3* carriers: • 280% higher glyburide plasma levels • 267% higher glimepiride plasma levels	[37]
			23 healthy volunteers: • Single 500 mg oral dose tolbutamide • Then 75 g oral glucose at 1, 4.5, and 8 hours after	Tolbutamide oral clearance	*CYP2C9*3/*3* individuals: • 6.5-fold lower tolbutamide oral clearance	[38]
			21 healthy volunteers: • Single 3.5 mg oral dose glyburide • Then 75 g glucose at 1, 4.5, and 8 hours after	Glyburide oral clearance	*CYP2C9*3/*3* individuals: • <50% lower glyburide oral clearance	[39]
			15 healthy volunteers: • Single 500 mg oral dose tolbutamide	• Concentration • Oral clearance	*CYP2C9*1/*2* individuals: • 1.5-fold increased tolbutamide concentration • 29%, greater reduction in tolbutamide oral clearance • 38%, greater reduction in tolbutamide formation clearance *CYP2C9*1/*3* individuals: • 1.9-fold increased tolbutamide concentration • 48%, greater reduction in tolbutamide oral clearance • 56%, greater reduction in tolbutamide formation clearance	[40]
			• T2DM patients • Tolbutamide-treated participants of Rotterdam study	• Fasting serum glucose levels • Required doses	*CYP2C9*3* carriers: • larger decrease in fasting serum glucose levels after therapy start • required lower doses	[42]
			Go-DARTs study: • 1073 sulfonylurea-treated T2DM patients	• HbA1C target • Glycemic goal achievement	*CYP2C9*2/2, CYP2C9*2/*3, CYP2C9*3/*3* individuals: • 3.4 times more likely to achieve HbA1C <7% • achieved glycemic goal more rapidly and persistently	[43]

TABLE 30.2 Continued

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
			207 sulfonylurea-treated T2DM patients	Mean sulfonylurea stable dose	*CYP2C9*3* carriers: • trended to lower mean stable glimepiride dose	[44]
		*CYP2C9*2* and **3*: • Compared with *CYP2C9*1/*1* (known *CYP2C19* genotype)	Clinical trial NCT0080601318: • 18 healthy Chinese volunteers: • Single dose 5 mg glipizide	Glipizide plasma concentrations over 36 hours	Compared with *CYP2C9*1/*1* individuals, *CYP2C9*1/*3* individuals: • higher glipizide AUC (by 95.5%) • lower glipizide clearance (−51.1%)	[41]
		*CYP2C9*2* and **3:* • Compared with CYP2C9*1/*1	80 glibenclamide-treated T2DM patients	• Response to glibenclamide • Glibenclamide plasma levels	*CYP2C9*1/*2* or **1/*3* patients: • more likely to have controlled T2DM (fasting plasma glucose value <110 mg/dl) compared with *CYP2C9*1/*1* patients *CYP2C9*2* and **3* carriers: • relatively higher plasma levels of glibenclamide compared with *CYP2C9*1/*1* patients	[45]
			• 20 T2DM patients with severe hypoglycemia • 337 T2DM without history of severe hypoglycemia	Severe hypoglycemia	*CYP2C9*3/*3* and **2/*3* genotypes: • increased risk of severe hypoglycemia (OR 5.2)	[47]
			• 92 T2DM with mild hypoglycemia • 84 T2DM patients without hypoglycemia	Mild hypoglycemia	*CYP2C9*3* allele: • increased risk of mild hypoglycemia (OR 1.7)	[48]
			• 102 T2DM patients with severe hypoglycemia • 101 T2DM without history of severe hypoglycemia	Severe hypoglycemia	• No over-representation of CYP2C9 PM phenotypes (*CYP2C9*2/*2, *2/*3*, or **3/*3* genotypes) in case group • Mean sulfonylurea doses in CYP2C9 EM *CYP2C9*1/*1, *1/*2*), IM (*CYP2C9*1/*3*), PM (*CYP2C9*2/*2, *2/*3, *3/*3*): • significantly lower in nonhypoglycemic controls compared with hypoglycemic cases	[50]

(Continued)

TABLE 30.2 Continued

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
			• 15 T2DM patients with mild hypoglycemia • 93 T2DM patients without hypoglycemia	Mild hypoglycemia	_CYP2C9_ genotypes with 1 (_*1/*2_or _*1/*3_) or 2 (_*2/*2_, _*2/*3_, or _*3/*3_) variant _CYP2C9_ alleles: • more frequent among patients reporting hypoglycemic attacks than among those not reporting (60 and 7% versus 39 and 3%)	[49]
Repaglinide	_CYP2C8_	_CYP2C8 *3_ compared with _CYP2C8*1/*1_	28 healthy volunteers administered 0.25 mg repaglinide	• Plasma concentration • Blood glucose levels 7 hours after dosing	_CYP2C8*1/*3_ individuals: • 45% lower repaglinide mean AUC and 39% lower peak plasma concentration	[65]
		• _CYP2C8 *3_ compared with _CYP2C8*1/*1_ • _SLCO1B1 −11187G>A_ • and _521T>C_	56 healthy volunteers administered 0.25 mg repaglinide	• Plasma concentration of repaglinide • Blood glucose levels 7 hours after dosing	_CYP2C8*1/*3_ individuals: • 48% lower repaglinide mean AUC • 44% lower peak plasma concentration _SLCO1B1 521CC_ individuals: • 107% and 188% higher AUC compared with _SLCO1B1 521TC_ and _TT_ individuals _CYP2C8*3_ allele and _SLCO1B1 521CC_ genotype: • independent predictors of AUC and concentration	[66]
Nateglinide	_CYP2C9_	• _CYP2C9*2_ • _CYP2C9*3_	Prospective clinical study: • 26 healthy volunteers receiving 180 mg nateglinide dose	Nateglinide clearance	_CYP2C9*3_ allele carriers: • significantly reduced oral nateglinide clearance	[67]
			Randomized, open-label, two-period crossover study: • 18 healthy subjects: 120 mg nateglinide dose alone/combination with CYP2C9 inhibitor sulfinpyrazone (start on day 7, 200 mg twice daily)	Nateglinide clearance	Co-administration of nateglinide with sulfinpyrazone: • 28% higher mean AUC	[68]
			35 Chinese male healthy volunteers: • Single 120 mg oral dose nateglinide	• Plasma concentrations • Blood glucose level	• _CYP2C9*3_ allele associated with nateglinide AUC and clearance • No significant association with changes in blood glucose-lowering effect	[69]

TABLE 30.2 Continued

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
Pioglitazone	*CYP2C8*	*CYP2C8*2* and **3*: • Compared with *CYP2C8*1/*1*	16 healthy volunteers: • Single 15 mg oral dose pioglitazone while treated either with placebo or 160 mg tripethoprim twice daily	Pioglitazone pharmacokinetics	• Pioglitazone AUC 34% smaller in *CYP2C8*3/*3* individuals and 26% smaller in *CYP2C8*1/*3* individuals than in *CYP2C8*1/*1* individuals • *CYP2C8* genotype not associated with plasma concentrations increase during trimethoprim phase	[79]
			31 healthy volunteers: • Receiving single 15 mg dose pioglitazone	Pioglitazone clearance	*CYP2C8*3* carriers: • significant effect on pioglitazone central compartment clearance	[80]
			30 healthy Caucasian volunteers: • Single 15 mg dose pioglitazone or 600 mg gemfibrozil	Pioglitazone clearance	*CYP2C8*3* carriers: • 29.7% lower mean AUC compared to *CYP2C8*1/*1* individuals when pioglitazone administered alone • 5.2-fold mean increase in AUC compared with 3.3-fold mean increase in *CYP2C8*1/*1* carriers when drug co-administered with gemfibrozil	[81]
			17 Healthy African American volunteers	Pioglitazone clearance	Keto-pioglitazone metabolite AUC ratio significantly lower in *CYP2C8*2* carriers than in *CYP2C8*1/*1* carriers	[82]
Rosiglitazone	*CYP2C9*	*CYP2C8*2* and **3*: • Compared with *CYP2C8*1/*1*	Randomized crossover study: • 8 healthy individuals in	Rosiglitazone AUC	• No effect on rosiglitazone metabolism • Trend toward increased rosiglitazone AUC in *CYP2C8*1/*1* individuals	[84]
			2-phase, open-label crossover trial: • 23 healthy volunteers	Rosiglitazone pharmacokinetics	No association with rosiglitazone pharmacokinetics	[85]
			31 healthy volunteers: • 8 mg rosiglitazone	• Rosiglitazone clearance • Elimination half-life • Glucose-lowering effect	• Mean total clearance higher in *CYP2C8*3* carriers • Elimination half-lives lower in *CYP2C8*1/*3* and *CYP2C8*3/*3* than in *CYP2C8*1/*1* carriers • No association of *CYP2C8* genotypes with rosiglitazone glucose-lowering effect	[86]

(Continued)

TABLE 30.2 Continued

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
		*CYP2C8*2* and **3*: • Compared with *CYP2C8*1/*1*	26 healthy volunteers	Rosiglitazone AUC and clearance	Compared with *CYP2C8*1/*1* individuals, *CYP2C8*1/*3* genotype: • significantly lower rosiglitazone AUC • significantly higher rosiglitazone oral clearance	[87]
			South Danish Diabetes Study: • 187 rosiglitazone-treated T2DM patients	• Rosiglitazone trough plasma concentration • Effect on HbA1c • Effect on risk of adverse events	*CYP2C8*3* carriers: • lower rosiglitazone plasma levels • reduced therapeutic response • lower edema risk	[88]
		*CYP2C8*11:* • Compared with *CYP2C8*1/*1*	9 healthy Asian volunteers: • 4 mg single dose rosiglitazone	Rosiglitazone plasma concentration	*CYP2C8*1/*11* carriers: • AUC/peak plasma concentration 54 and 34% higher than mean values in control *CYP2C8*1/*1* individuals	[89]

Note: These polymorphisms are associated with oral antidiabetic drug metabolism, drug response, and incidence of adverse reactions.

TABLE 30.3 Gene Polymorphisms of Transporters and Receptors

Drug	Gene (Protein)	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Refs
Metformin	*SLC22A1* (OCT1)	• *R61C,* • *G401S* • *M420del* • *G465R*	20 healthy volunteers: • two subsequent metformin doses (1000 mg evening, 850 mg next morning)	Metformin plasma and urine concentrations	Defective variants associated with significantly decreased glucose-lowering response	[17]
			20 healthy volunteers: • two subsequent metformin doses (1000 mg, then 850 mg after 12 hours)	Metformin pharmacokinetics	OCT1 genotypes carrying reduced-function allele: • significant effect on metformin pharmacokinetics associated with ◦ higher AUC ◦ maximum plasma concentration ◦ lower oral distribution volume	[18]
			103 healthy volunteers	Metformin renal clearance	OCT1 expression associated with: • decreased hepatic uptake • increased renal clearance • metformin response	[19]
		11 tagging SNPs	102 T2DM patients on metformin	HbA1C change	Association of rs622342 *A>C* with HbA1C reduction: • for each rs622342 *C* allele, reduction in HbA1C was 0.28% less	[20]

TABLE 30.3 Continued

Drug	Gene (Protein)	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Refs
		• *R61C* • *G401S* • *M420del* • *G465R*	138 T2DM patients randomized to: • metformin 1000 mg/day for 4 weeks • 2000 mg/day after	Metformin response	Reduced-function diplotypes: • *R61C, G401S, M420del*, and *G465R* in OCT1 associated with ◦ significant decrease in trough metformin concentration ◦ reduced absolute decrease in HbA1c	[21]
		P341L	96 healthy male volunteers: • single 500 mg dose metformin	Metformin pharmacokinetics	*OCT1 P341L* associated with: • increased AUC • altered metformin pharmacokinetics	[22]
		• *R61C* • *G465R*	246 T2DM patients: • metformin treatment • 53 common gastrointestinal side effects	Metformin gastrointestinal adverse effects	OCT1 *R61C* and *G465R* alleles: • metformin intolerance	[23]
	SLC22A2 (OCT2)	*808G>T*	15 healthy volunteers: • single 500 mg oral dose	Metformin renal tubular clearance	Associated with metformin renal tubular clearance	[26]
		808G>T	23 healthy volunteers: • single 850 mg oral dose	Metformin pharmacokinetics	Compared with reference *808GG* carriers, *808T* carrier: • significantly higher mean renal clearance	[27]
	SLC22A2 (OCT2)	• *596C>T* • *602C>T* • *808G>T*	Healthy volunteers	Metformin pharmacokinetics	Compared with reference genotype, variant allele carriers: • higher peak plasma concentration • higher AUC • lower renal clearance	[25]
	• *SLC22A1* (OCT1) • *SLC22A2* (OCT2)	• OCT1 *-43T>C* • OCT2 *Val408Met*	33 T2DM patients on metformin: • 24 responders • 9 non-responders	Metformin clinical efficacy	Little contribution to metformin clinical efficacy	[15]
		• OCT1 rs622342 • OCT2 rs316019	148 T2DM patients	Reduction of HbA1c	No association with metformin response	[24]
	SLC47A1 (MATE1)	rs2289669 *G>A*	116 T2DM patients on metformin	HbA1C reduction	For each MATE1 rs2289669 *A* allele: • HbA1C reduction of 0.30% more • interaction between OCT1 rs622342 *A>C* and MATE1 rs2289669 *G>A* with respect to HbA1C reduction	[29]

(Continued)

TABLE 30.3 Continued

Drug	Gene (Protein)	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Refs
		rs8065082 (in tight LD with rs2289669 *G>A*)	Diabetes Prevention Program: • 2994 individuals • 990 metformin-treated	Response to metformin	MATE1 rs8065082: • suboptimal responses to metformin	[31]

	• *SLC47A1* (MATE1)
	• *SLC22A1* (OCT1)
		• MATE1 rs2289669 *G>A*
		• OCT1 rs622342 *A>C*
			Rotterdam study:
			• 98 T2DM patients on metformin
				HbA1C reduction	• Interaction between OCT1 rs622342 *A>C* and MATE1 rs2289669 *G>A* with respect to HbA1C reduction
					• Decrease in HbA1C levels larger for each OCT1 rs622342 *A* allele and MATE1 rs2289669 *A* allele
					• *AA* homozygotes for both genes:
					  • average decrease in HbA1C largest (−0.91%)
					  • in OCT1 rs622342 CC and MATE1 rs2289669 GG participants
					  • HbA1C levels increased 0.48%
						[30]

Sulfonylureas	*ABCC8*	*S1369A*	115 T2DM gliclazide-treated patients	HbA1C decrease	Compared to noncarriers, *ABCC8 1369A* carriers: • more significant HbA1C decrease after treatment (−1.6% vs −0.76%)	[55]
	• *ABCC8* • *KCNJ11*	• *ABCC8 S1369A* • *KCNJ11 E23K*	1268 Chinese T2DM patients: • gliclazide-treated 8 weeks	• Fasting plasma glucose decrease • 2-hour glucose after oral glucose tolerance test • HbA1C decrease	*KCNJ11 23K*, and *ABCC8 1369A* carriers: • greater decreases in fasting plasma glucose *ABCC8 1369SA* and *1369AA* T2DM patients: • greater decreases in fasting plasma glucose (−2.8% and −7.7%, respectively) compared to *ABCC8 1369SS* carriers • greater decreases in 2-hour glucose after oral glucose tolerance test (−10.8% and −11.9%, respectively) compared to *ABCC8 1369SS* patients • greater decreases in HbA1C from baseline (−1.5% and −1.7%, respectively) compared to −1.2% *ABCC8 1369SS* patients *ABCC8 1369SA* and *1369AA* patients: • greater chance of responding to gliclazide therapy (OR 1.4 and 2.2, respectively) compared to *ABCC8 1369SS* patients	[56]

TABLE 30.3 Continued

Drug	Gene (Protein)	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Refs
	ABCC8	*S1369A*	• 32 T2DM patients with severe hypoglycemia • 125 T2DM without history of severe hypoglycemia	Severe hypoglycemia	*A* variant of *S1369A* polymorphism in *ABCC8*: • lower risk for severe hypoglycemia	[60]
	KCNJ11	*E23K*	101 T2DM patients: • sulfonylureas as add-on after failure to achieve glycemic goals while on metformin monotherapy	HbA1C decrease	Compared with *EE* carriers, *KCNJ11 23K* carriers: • greater decrease in HbA1c after 6 months treatment	[57]
			100 Egyptian T2DM patients	• Sulfonylurea response • Secondary failure	*KCNJ11 23K*: • secondary sulfonylurea failure	[58]
		E23K	• 43 T2DM patients with severe hypoglycemia • 54 T2DM without history of severe hypoglycemia	Severe hypoglycemia	• No association with severe hypoglycemia *K* variant of *E23K* in *KCNJ11*: • reduced response to sulfonylurea therapy and lower risk of severe hypoglycemia	[61]
		E23K	• 92 T2DM patients with mild hypoglycemia • 84 T2DM without hypoglycemia	Mild hypoglycemia	No association with mild hypoglycemia	[59]
Repaglinide	*SLCO1B1*	*CYP2C8 *3* compared with: • *CYP2C8*1/*1* • *SLCO1B1 −11187G>A* • *521T>C*	56 healthy volunteers: • 0.25 mg repaglinide	• Plasma concentration of repaglinide • Blood glucose levels 7 hours after dosing	*CYP2C8*1/*3* individuals: • 48% lower repaglinide mean AUC and 44% lower peak plasma concentration *SLCO1B1 521CC* individuals: • 107% and 188% higher repaglinide AUC compared with *SLCO1B1 521TC* and *TT* individuals *CYP2C8*3* allele and *SLCO1B1 521CC* genotype: • independent predictors of repaglinide AUC and concentration	[66]

(Continued)

TABLE 30.3 Continued

Drug	Gene (Protein)	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Refs
		SLCO1B1 521T>C	28 healthy volunteers: • single 0.25, 0.5, 1, or 2 mg repaglinide dose	Repaglinide metabolites plasma concentrations	*SLCO1B1 521CC* individuals: • higher AUC compared with *SLCO1B1 521TT* individuals *SLCO1B1 521CC* individuals: • 60–70% higher max concentration compared with *SLCO1B1 521TT* individuals • same tendency between genotypes toward greater lowering of blood glucose concentrations	[72]
		SLCO1B1 521T>C	32 _CYP2C8*1/*1_ healthy: • single 0.5 mg dose repaglinide	Repaglinide pharmacokinetics	*SLCO1B1 521CC* individuals: • 59% and 72% higher AUC compared with *SLCO1B1 521TC* and *TT* individuals *SLCO1B1 521CC* individuals: • 27% and 62% higher max repaglinide concentration compared with *SLCO1B1 521TC* and *TT* individuals • same tendency between genotypes toward greater lowering of blood glucose concentrations	[73]
		SLCO1B1 521T>C	22 healthy male Chinese individuals: • 2 mg repaglinide	Repaglinide pharmacokinetics and pharmacodynamics	• 521C allele attenuated repaglinide clearance • no significant differences in blood glucose-lowering effect observed	[74]
	KCNJ11/ ABCC8	• *KCNJ11 E23K* • *ABCC8* exon *16–3*	101 Chinese T2DM patients • 24-week repaglinide	Repaglinide efficacy	*KCNJ11 23EK* and *23KK* carriers: • higher HbA1C decrease *ABCC8* exon 16–*3CC* homozygous genotype: • better insulin sensitivity	[76]
	KCNJ11	*E23K*	40 randomly selected Chinese T2DM patients: • 8-week repaglinide treatment regimen	Repaglinide efficacy	*23E* allele carriers had higher HbA1c levels compared to 23KK carriers	[77]
Nateglinide	*SLCO1B1*	*SLCO1B1 521T>C*	17 healthy volunteers: • 90 mg single dose nateglinide	Nateglinide plasma concentrations	Compared with *SLCO1B1 521TT* carriers, *SLCO1B1 521CC* carriers: • higher maximum concentration and AUC (76% versus 108%) Compared with *SLCO1B1 521TT* carriers, *SLCO1B1 521TC* carriers: • higher maximum concentration and AUC (82% versus 83%)	[75]

TABLE 30.3 Continued

Drug	Gene (Protein)	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Refs
			32 *CYP2C9*1/*1* healthy volunteers: • single 60 mg dose nateglinide	Nateglinide pharmacokinetics	No association with nateglinide disposition	[73]
			35 healthy Chinese male volunteers: • different *CYP2C9* and *SLCO1B1* genotypes • single 120 mg oral dose nateglinide	Nateglinide metabolism	*SLCO1B1 521C* allele alone: • no effect on nateglinide pharmacokinetics *CYP2C9*3* and *521 T>C* SNPs: • significant predictors of AUC and clearance • no association with blood glucose-lowering effect	[69]
Pioglitazone	*SLCO1B1*	*SLCO1B1 521T>C*	32 healthy volunteers: • 15 mg pioglitazone	Pioglitazone pharmacokinetics	No association with pioglitazone pharmacokinetic variables	[91]
	PPARγ	*PPARγ Pro-12Ala*	83 obese postmenopausal women: • 15 mg pioglitazone /day, 15 days	Metabolic and hormonal response to pioglitazone	• *PPARγ 12ProAla* genotype associated with greater decrease in serum glucose levels • Obese women carrying *12Ala* allele may benefit from pioglitazone treatment	[93]
		• *PPARγ Pro-12Ala* • *PPARγ Gly482Ser*	250 T2DM patients: • 30 mg pioglitazone, 24 weeks	Therapeutic response	*12Ala* allele: • more frequent in pioglitazone responders • greater decrease in HbA1C and fasting glucose plasma levels after 24 weeks No association between *Gly482Ser* and investigated parameters	[94]
		PPARγ Pro-12Ala	30 T2DM patients • pioglitazone monotherapy	• Glycemic target • Primary pioglitazone failure	No association with investigated parameters	[97]
		PPARγ Pro-12Ala	101 T2DM patients: • 15 mg/day pioglitazone for 12 weeks	Pioglitazone response	Compared to *12ProPro* carriers, *12Ala* carriers: • better therapeutic response • difference between groups not statistically significant	[95]
			67 T2DM patients: • 30 mg/day pioglitazone for 3 months	• Fasting plasma glucose • Postprandial plasma glucose • HbA1c • Lipid profile	Compared to *12ProPro* carriers, *12Ala* carriers: • significantly higher differential postprandial plasma glucose • significantly higher serum triglyceride values	[96]

(Continued)

TABLE 30.3 Continued

Drug	Gene (Protein)	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Refs
Rosiglitazone	_SLCO1B1_	_CYP2C8*2_ and _*3_ compared with • _CYP2C8*1/*1_ • _SLCO1B1 521T>C, −11187G>A, −10499A>C_ • _388A>G_	26 healthy volunteers	Rosiglitazone AUC and clearance	Compared with _CYP2C8*1/*1_ carriers, _CYP2C8*1/*3_ carriers: • significantly lower rosiglitazone AUC • significantly higher rosiglitazone oral clearance No association between _SLCO1B1_ SNPs and investigated parameters	[87]
		SLCO1B1 521T>C	32 healthy volunteers: • 4 mg rosiglitazone	Rosiglitazone pharmacokinetics	No association with rosiglitazone pharmacokinetics	[91]
	PPARγ	_PPARγ2 Pro-12Ala_	198 T2DM patients: • 4 mg/day rosiglitazone, 12 weeks	Decrease in fasting plasma glucose levels and HbA1C	_12Ala_ allele: • greater decrease in fasting plasma glucose levels • HbA1C decrease • better therapeutic response	[98]
Troglitazone	_PPARγ_	_PPARγ_ screening	TRIPOD study: 93 nondiabetic women: • previous gestational diabetes	Troglitazone response	_PPARγ_ variation underlies response to thiazolidinediones	[99]
		PPARγ Pro-12Ala	DPP study: 554 nondiabetic patients • gestational diabetes randomized to troglitazone	Progression from impaired glucose tolerance to diabetes	• _12Ala_ allele protected from diabetes development • _12ProPro_ carriers more likely to develop diabetes (OR 1.24)	[100]

Note: These polymorphisms are associated with oral antidiabetic drug response and incidence of adverse reactions.

TABLE 30.4 Gene Polymorphisms in Underlying Disease Condition or Insulin Secretion Pathway Associated with Oral Antidiabetic Drug Response

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
Metformin	_ATM_	rs112112617_C>A_	• GWAS in 1024 Scottish T2DM individuals • 2 replication cohorts in 1783 Scottish individuals • 113 individuals from UKPDS	Metformin treatment success	_ATM_ SNP rs112112617 _C_: • treatment success in combined all-patient meta-analysis	[32]

TABLE 30.4 Continued

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
			Metformin-treated T2DM patients: • DCS West-Friesland study (n = 929) • Rotterdam Study (n = 182) • UK CARDS (n = 254) • 5-cohort meta-analysis	Metformin effect on delaying disease progression to diabetes caused by impaired glucose tolerance	• *ATM* rs112112617 *C* allele associated with metformin response in DCS cohort • Similar but insignificant trend in Rotterdam Study • No association in CARDS cohort • 3-cohort meta-analysis: ◦ *ATM* rs112112617 *C* allele associated with increased metformin response in whole patient group (OR 1.24, p = 0.016) ◦ likewise for T2DM patients starting metformin monotherapy (OR 1.38, p = 0.015) • Combined 5-cohort meta-analysis: ◦ C allele associated with −0.050 HbA1c decrease per allele copy	[33]
			DPP: 988 T2DM patient participants on metformin	Metformin response	No association with metformin response	[34]
Sulfonylureas		• *TCF7L2* rs7903146 (*C*>*T*) • rs12255372 (*G*>*T*) polymorphisms	Go-DARTs: 901 sulfonylurea-treated T2DM patients	• Treatment failure • HbA1C target achievement	• More treatment failure observed in *TT* homozygotes of either SNP • 57% *TT* homozygotes failed to reach HbA1C target <7% compared with 40% *GG* individuals • Heterozygote group: intermediate failure rate • OR treatment failure per analyzed allele overall: ◦ 1.27 for rs7903146 ◦ 1.28 for rs12255372 • Both polymorphisms influenced initial sulfonylurea success: ◦ rs12255372—majority of patients carrying 2 *T* allele copies (57%) not likely to achieve HbA1C<7% in 1 year of treatment	[62]

(Continued)

TABLE 30.4 Continued

Drug	Gene	Polymorphisms	Subject Population	Investigated Parameter	Primary Outcome	Reference
		TCF7L2 rs7903146 (*C>T*)	87 T2DM patients 6 months on sulfonylurea plus metformin	• HbA1C reduction • Fasting plasma glucose	rs7903146 *T* carriers (genotypes *CT, TT*): • significantly lower HbA1C and fasting plasma glucose reduction after 6-month treatment	[63]
		• *TCF7L2* rs7903146 (*C>T*)	189 T2DM patients treated with sulfonylureas	Sulfonylurea secondary failure	*TCF7L2* genotype only predictor of treatment failure (OR 1.57)	[64]

TABLE 30.5 Clinical Trials Assessing Oral Antidiabetic Drug Pharmacogenomics

Drug	Polymorphisms	Clinical Trial	Status	Population	Outcome	Reference
Metformin	OCT1 • *R61C* • *G401S* • *M420del* • *G465R*	NCT00187681	Completed	226 healthy volunteers: • 2 doses metformin (1000 mg evening; 850 mg next morning)	• OCT1-defective variants: significantly decreased glucose-lowering response • Higher AUC • Maximum plasma concentration • Lower oral distribution volume	[17,18]
Metformin	• OCT • OCT2	NCT01341795	Unknown	T2DM patients	Metformin response	NA
Metformin	OCT2 *A270S*	NCT01237522	Completed	Healthy Caucasian volunteers	Metformin renal clearance after single 500 mg dose	NA
Metformin	OCT2 808G>T	NCT00187720	Completed	23 healthy volunteers: • single 850 mg oral dose	*808T* carriers: significantly higher mean renal clearance compared with *808GG* carriers	[27]
Metformin	Genotypic determinants of metformin response	NCT00780715	Recruiting	• T2DM drug-naïve patients treated with 500 mg metformin • T2DM patients who failed on metformin to be randomized to other antidiabetic medications	• Metformin response • Patient monitoring: 6 months	NA
Metformin	• OCT • MATE1 • MATE2-k	NCT01681680	Not yet recruiting	Healthy Caucasian and Asian volunteers	• Metformin pharmacokinetics • Metformin glucose-lowering response	NA

TABLE 30.5 Continued

Drug	Polymorphisms	Clinical Trial	Status	Population	Outcome	Reference
Sulfonylureas (Glibenclamide)	*CYP2C9* gene polymorphisms	NCT01242137	Unknown	T2DM patients	Sulfonylurea pharmacokinetics	NA
Sulfonylureas (Glipizide)	*CYP2C9* polymorphisms with known *CYP2C19* genotype	NCT00806013	Completed	18 healthy Chinese individuals: • single 5 mg dose glipizide • plasma concentrations determined for 36 hours	Compared with *CYP2C9*1/*1* carriers, *CYP2C9*1/*3* carriers: • higher glipizide AUC (by 95.5%) • lower gliclazide clearance (by 51.1%)	[41]
Sulfonylureas (Tolbutamide)	*CYP2C9*	NCT01061112	Recruiting	Individuals from pharmacogenetics registry with known drug metabolism genotype	Extent of CYP2C9 inhibition in tolbutamide-treated: • *CYP2C9*1/*1* carriers • *CYP2C9*1/*3* carriers • *CYP2C9*3/*3* carriers	NA
Sulfonylureas	*TCF7L2*	NCT00738088	Ongoing	T2DM patients genetically predisposed to diabetes due to changes in *TCF7L2*	Response to sulfonylurea medication	NA
Pioglitazone	Genetic predictors of pioglitazone response	NCT01135394	Recruiting	T2DM patients	Pioglitazone response	NA

Source: From www.clinicaltrials.gov/.

In another study, Becker et al. analyzed the association of 11 tagging SNPs of the OCT1 with HbA1c change in 102 metformin users. They found significant association of the rs622342*A>C* polymorphism with HbA1c reduction [20]. The association of OCT1 defective variants with metformin response has been replicated in three more recent pharmacogenomic studies. In 138 T2DM patients randomized to metformin 1000 mg/day for 4 weeks and afterward to 2000 mg/day, reduced-function diplotypes composed of the alleles *R61C*, *G401S*, *M420del*, and *G465R* in OCT1 were associated with a significant decrease in trough metformin steady-state concentration and a reduced absolute decrease in HbA1c [21].

A study of 96 healthy Korean male volunteers administered a single 500 mg dose of metformin found that the OCT1 *P341L* polymorphism was associated with increased AUC and altered metformin pharmacokinetics [22]. Beyond the association of OCT1 variants with metformin pharmacokinetics and drug response, the findings of a recent study of 246 T2DM patients on metformin treatment, of whom 53 experienced common gastrointestinal side effects, showed an association of two OCT1 defective alleles (*R61C* and *G465R*) with metformin intolerance [23]. The researchers suggested that metformin intolerance is induced by a local increase in metformin concentration in the intestinal tissue and that, since OCTs are also expressed in intestinal cells, altered OCT expression potentially mediates the experience of metformin-induced adverse events.

OCT1 gene polymorphisms are associated with metformin response with a high degree of reproducibility and may emerge as useful genetic markers in personalizing metformin treatment. However, currently only one clinical trial assessing the effect of OCT1 genetic variants on metformin response in T2DM patients (Table 30.5) is registered on the ClinicalTrials.gov website (www.clinicaltrials.gov/).

Results from prospective, randomized clinical trials on the association of OCT1 gene polymorphisms with metformin response will enable the clinical application of metformin pharmacogenomics in routine clinical practice.

30.5.1.2 OCT2: SLC22A2 Gene Polymorphisms and Metformin Response

The *SLC22A2* gene encodes for OCT2 and is located on chromosome 6q25.3; it consists of 11 exons spanning 42 kb. The human *SLC22A2* gene is highly polymorphic, and numerous polymorphisms have been described in various populations (www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=6582). The most common *OCT2* polymorphisms, their MAF, and their effect on transporter activity are presented in Table 30.1.

In several studies, OCT2 gene polymorphisms were analyzed in association with metformin response together with OCT1; however, no significant results [19,21,23,24] and little contribution to the clinical efficacy of metformin [15] were derived. Nevertheless, evidence on the association of OCT2 gene polymorphisms with metformin pharmacokinetics first appeared in studies on healthy volunteers, suggesting that a decrease in transport function associated with OCT2 variants results in reduced renal clearance of metformin and consequently leads to increased plasma concentrations. More specifically, in healthy participants, Song et al. showed that OCT2 *596C>T*, *602C>T*, and *808G>T* variants were associated with increased metformin AUC and decreased renal clearance [25], whereas Wang et al., demonstrated that the *808G>T* OCT2 gene polymorphism was associated with metformin renal tubular clearance [26]. In another study, controversial results were obtained, since the same variant (OCT2 *808G>T*) was correlated with higher mean renal clearance compared with individuals with the reference genotype (*808GG*) [27].

Four ongoing clinical trials assessing the effect of OCT2 gene polymorphisms on metformin response are registered on the ClinicalTrials.gov website (listed in Table 30.5). Clinical trial NCT01237522, assessing the effect of the OCT2 *A270S* polymorphism on metformin renal clearance, has been completed; however, the results have not yet been published. The effect of OCT2 genetic variants on metformin pharmacokinetics and response is currently being investigated in three clinical trials (NCT00780715, NCT01681680, and NCT01341795). As in the case of OCT1, it is expected that the results of ongoing clinical trials will shed light on metformin's pharmacogenomics.

30.5.1.3 MATE1: SLC47A1 Gene Polymorphisms and Metformin Response

The human MATE1 gene (*SLC47A1*) is located on chromosome 17p11.2 and encodes a protein of 570 amino acids [28]. MATE1 is mainly expressed in the liver and kidney and has an important role in the renal excretion of metformin. More than 70 polymorphisms have been identified in *SLC47A1* (www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=55244); however, only a few of them appear to have an association with metformin clearance and response (Table 30.1).

Results from studies with metformin-treated T2DM patients show a consistent association of two highly linked MATE1 variants (rs2289669*G>A* and rs8065082*C>T*) with HbA1c reduction and metformin response. In a study of 116 metformin-treated T2DM patients, each MATE1 rs2289669 *A* allele was associated with 0.30% increased reduction of HbA1c [29]. The same research group looked at whether there is an interaction between the OCT1 rs622342 *A>C* and the MATE1 rs2289669 *G>A* polymorphisms with respect to HbA1c reduction in 98 metformin users who were participants in the Rotterdam Study [30]. In accordance with previous results, each OCT1 rs622342 *A* allele and MATE1 rs2289669 *A* allele was associated with greater decrease in HbA1c levels. Especially in *AA* homozygotes for both genes, the average decrease in HbA1c was the largest (−0.91%), whereas in OCT1 rs622342 *CC* and MATE1 rs2289669 *GG* participants, HbA1c levels increased by 0.48% (Table 30.3) [30].

In another large study, involving 990 metformin-treated individuals who participated in the Diabetes Prevention Program (DPP), the MATE1 rs8065082 gene polymorphism, which is in tight LD with MATE1 rs2289669, was associated with suboptimal responses to metformin treatment [31]. More recently, the MATE1 rs2289669 *A* allele was associated with higher HbA1c reduction in 148 metformin-treated T2DM patients [24]. However, the association of MATE1 gene polymorphisms with metformin response was not replicated in all studies [21]. The association of *MATE1* genetic variations with metformin response is the subject of clinical trial NCT01681680 (Table 30.5), which will assess the effect of OCT2, MATE1, and MATE2-K genetic variations with metformin pharmacokinetics and metformin glucose-lowering response in healthy Caucasian and Asian volunteers.

30.5.1.4 Ataxia Telangiectasia Mutated: Gene Polymorphisms and Metformin Response

Ataxia telangiectasia mutated (*ATM*) is an important tumor suppressor gene with key roles in DNA repair and cell cycle control. *ATM* gene SNP rs11212617*C>A* was identified in a large GWAS as a significant mediator of metformin response. This GWAS, jointly conducted by the Genetics of Diabetes Audit and Research Tayside (GoDARTs) study group and the UK Prospective Diabetes Study (UKPDS) pharmacogenetics study group, included 1024 Scottish T2DM individuals (Table 30.4). Results were replicated in two cohorts that included 1783 Scottish individuals and 113 individuals from the UKPDS. In the combined meta-analysis of all patients, the *ATM* SNP rs11212617 *C* allele was associated with metformin treatment success, both in terms of achieving Hb1A1c values $\leq 7\%$ and in terms of a lowering of HbA1c when analyzed as a quantitative trait [32]. ATM is estimated to alter response to metformin through an upstream effect of AMP-activated protein kinase [32]. The UK study group attempted to replicate their original finding in three independent cohorts using a phenotypic definition identical to that used in the original discovery [33].

The UKPDS researchers included metformin-treated T2DM patients from Diabetes Care System West-Friesland (DCS), from the Rotterdam Study from the Netherlands, and from the multicenter UK Collaborative Atorvastatin Diabetes Study (CARDS) trial. Data were analyzed in the individual cohorts, separately, combined in a meta-analysis, and combined in an additional meta-analysis that included the GoDARTS and UKPDS stage 2 replication cohorts. According to UKPDS results, the original association of the *ATM* rs112112617 *C* allele with metformin response was replicated in the DCS cohort; also, a similar but nonsignificant trend was present in the Rotterdam Study; no association was found in the CARDS cohort [33]. In the meta-analysis of the three cohorts, *ATM* rs11212617 *C* was associated with increased metformin response in the entire group of patients (OR 1.24; $p=0.016$), as well as in T2DM patients starting metformin monotherapy (OR 1.38; $p=0.015$). In the combined meta-analysis of all five replication cohorts, the *C* allele was associated with an HbA1c decrease of -0.050 per allele copy. When HbA1c was analyzed as a quantitative trait, no significant effect of the *ATM* gene SNP on metformin response was observed.

In a more recent study of 988 metformin-treated T2DM patient participants in the DPP, this association was not replicated because the rs11212617 *C* allele conferred no detectable advantage on metformin recipients [34]. However, among other reasons that may explain the controversy in the results, metformin response is defined differently in a prediabetic cohort such as the DPP; therefore, the association of the *ATM* rs11212617 gene polymorphism with metformin response merits additional follow-up in independent cohorts.

The association of *ATM* gene polymorphisms with metformin response is a quite promising pharmacogenetic link to personalized metformin therapy. Currently, no clinical trials are registered at ClinicalTrials.gov on *ATM* genetic variations and metformin response; however, it is anticipated that such trials will be initiated soon.

30.5.2 Sulfonylureas

Sulfonylurea oral hypoglycemic agents are the oldest class of oral diabetes medication and have been a mainstay of T2DM pharmacotherapy for over 50 years. They are relatively cost-effective and well tolerated. Sulfonylureas—of both the first (tolbutamide, tolazamide, chlorpropamide, and acetohexamide) and the second generation (glibenclamide or glyburide, glipizide, gliclazide, and glimepiride)—act by blocking ATP-dependent potassium (K_{ATP}) channels in pancreatic B cells thereby mimicking the effect of glucose on the production of a triggering signal for insulin secretion in a glucose-independent manner. Sulfonylureas are among the most potent oral antidiabetic medications, having an HbA1c-lowering efficacy of 1–2%. Their major adverse effects are hypoglycemia and weight gain at initiation of therapy [11]. According to the ADA, sulfonylurea treatment is included in well-validated core therapies for T2DM to achieve and maintain HbA1c levels lower than 7% [11]; however, American Association of Clinical Endocrinologists (AACE)/American College of Endocrinology (ACE) algorithms have assigned a lower priority to sulfonylureas because they increase the risk of hypoglycemia [35].

Significant interindividual variability exists among patients who are on sulfonylurea treatment in terms of HbA1c decrease, secondary failure of therapy, and incidence of hypoglycemic events. It is widely accepted that response to sulfonylureas is genetically determined by genetic polymorphisms affecting their pharmacokinetics and pharmacodynamics, as well as by polymorphisms in genes that are associated with T2DM incidence.

30.5.2.1 Sulfonylurea Pharmacokinetics

Cytochrome P450 2C9 (CYP2C9)

All sulfonylureas are extensively metabolized in the liver, primarily by the cytochrome P450 2C9 isoenzyme (CYP2C9) [36]. Polymorphisms in the *CYP2C9* gene seriously affect the enzymatic activity of encoded CYP2C9. Based on phenotype, the population can be divided into extensive, intermediate, and poor metabolizers (EMs, IMs, and PMs, respectively). To date, more than 50 different allelic variants have been identified in the *CYP2C9* gene (www.cypalleles.ki.se/cyp2c9.htm). Among them, the most common are **2* (rs1799853, *430C>T*, Arg144Cys) and **3* (rs1057910, *1075A>C*, Ile359Leu) (Table 30.1), which allow for prediction of more than 85% of PMs in most populations, including Caucasians [36].

Altered pharmacokinetics of sulfonylureas due to polymorphisms in the *CYP2C9* gene may partially explain the differences in patient dose requirements and frequency of adverse reactions. The effect of these polymorphisms on sulfonylurea pharmacokinetics has been investigated in several studies in healthy volunteers (Table 30.2). In principle, the *CYP2C9 *2* and **3* alleles are associated with impaired metabolism and reduced oral clearance of sulfonylureas. Plasma concentration of glyburide and glimepiride has been shown to be higher in *CYP2C9*3* allele carriers (280% and 267%, respectively) than in *CYP2C9*1/*1* carriers [37]. Similarly, *CYP2C9*3/*3* individuals had a 6.5-fold lower oral clearance of tolbutamide and less than 50% lower oral clearance of glyburide compared with *CYP2C9*1/*1* individuals [38,39].

In a study of 15 healthy volunteers (5 per genotype group *CYP2C9*1/*1*, **1/*2*, and **1/*3*) receiving a single oral dose of 500 mg tolbutamide, both plasma levels and oral clearance were elevated in **1/*2* and **1/*3* individuals compared to **1/*1* individuals [40]. More recently, the effect of *CYP2C* polymorphisms on the pharmacokinetics of glipizide was assessed (Table 30.5) in a clinical trial (NCT00806013). Eighteen healthy Chinese subjects were divided into three groups (*CYP2C9*1/*1* and CYP2C19 EMs; *CYP2C9*1/*1* and CYP2C19 PMs; and *CYP2C9*1/*3* and CYP2C19 EMs) and were administered a single dose of glipizide (5 mg). Drug plasma concentrations were determined during a 36-h period. The results showed that *CYP2C9*1/*3* individuals had higher glipizide AUC (by 95.5%) and lower glipizide clearance (−51.1%) compared with *CYP2C9*1/*1* individuals [41].

Subsequent results from studies on T2DM patients treated with sulfonylureas suggest that there is a straight linear effect of *CYP2C9* gene polymorphisms on sulfonylurea pharmacokinetics and therapeutic response, estimated by fasting serum glucose and HbA1c decrease, dose requirement, or incidence of hypoglycemic events. Specifically, among T2DM patients treated with tolbutamide, *CYP2C9*3* carriers had a greater decrease in fasting serum glucose levels after the start of therapy and required lower doses compared with *CYP2C9*1/*1* individuals [42]. In terms of achieving HbA1c goals, results derived from a Go-DARTs study of 1073 sulfonylurea-treated T2DM patients showed that carriers of two each of the *CYP2C9*2* and/or *CYP2C9*3* alleles were 3.4 times more likely to achieve HbA1c levels below 7%, compared with *CYP2C9*1/*1* patients [43]. Furthermore, these carriers also achieved glycemic goals more rapidly and persistently compared with *CYP2C9*1/*1* patients.

In a study on sulfonylurea-treated T2DM patients, *CYP2C9*3* carriers trended toward a lower stable glimepiride dose compared with *CYP2C9*1/*1* patients; however, this difference was not statistically significant [44]. In another, more recent study of 80 T2DM patients treated with glibenclamide, it was shown that those with a *CYP2C9*1/*2* or **1/*3* genotype were more likely to have controlled T2DM (fasting plasma glucose value < 110 mg/dl) compared with *CYP2C9*1/*1* patients [45]. Also, carriers of *CYP2C9*2* and **3* had relatively higher plasma levels of glibenclamide compared with *CYP2C9*1/*1* patients; however this difference was not statistically significant.

Dose titration of sulfonylureas in carriers of *CYP2C9* defective alleles can serve as an important tool for the application of pharmacogenomics and should also be taken into account in studies assessing the effect of *CYP2C9* genotypes on glycemic response. Taken together, the results of pharmacogenomic studies on the pharmacokinetics of sulfonylureas suggest that the presence of defective *CYP2C9* alleles is correlated with increased sulfonylurea plasma levels and may lead to an improved therapeutic effect of these drugs for *CYP2C9* defective allele carriers.

Hypoglycemia is the major and most common sulfonylurea adverse event, and can have a profound impact on T2DM patients' quality of life of [46]. Among other factors, increased plasma concentrations of sulfonylureas, decreased clearance, and prolonged action due to diminished drug metabolism in *CYP2C9* defective allele carriers might also be associated with hypoglycemia. The results of several studies so far seem to support this hypothesis. Specifically, it has been shown that carriers of *CYP2C9*3* were at higher risk of experiencing either severe (OR 5.2) [47] or mild hypoglycemic events (OR 1.7) [48,49].

In a recent study, Holstein et al. prospectively examined *CYP2C9*2* and **3* in association with severe hypoglycemia [50]. They did not confirm their initial results on the association of CYP2C9 PMs with hypoglycemia; however, when they calculated mean doses among different *CYP2C9* genotypes between hypoglycemic cases and nonhypoglycemic controls, they found that patients in the nonhypoglycemic control group were on significantly lower mean sulfonylurea dose compared with those in the hypoglycemic case group. The authors concluded that their results, despite the lack of association, point to the benefit of individualized and pharmacogenetic-based dosing of sulfonylureas to avoid incidence of hypoglycemic events, because the absence of association of *CYP2C9* polymorphisms with severe hypoglycemia in this study could be due to dose titration (lower dose) of sulfonylureas in nonhypoglycemic controls [50].

As mentioned earlier, AACE/ACE algorithms have moved sulfonylureas to lower priority because of their increased risk of hypoglycemia, despite their unquestionable therapeutic efficacy [35]. Therefore, pharmacogenetic markers such as *CYP2C9* could serve as useful tools for identifying T2DM patients who are at increased risk of sulfonylurea-induced hypoglycemia, and could increase the overall therapeutic value of these drugs.

Currently two clinical trials on the association of *CYP2C9* gene polymorphisms with sulfonylurea pharmacokinetics and response are registered at ClinicalTrials.gov (see Table 30.5). The clinical trial NCT01242137 will assess the dose requirements of T2DM patients treated with sulfonylurea (glibenclamide) based on their *CYP2C9* genotype. In this trial, investigators will test the hypothesis that patients with less active alleles of the CYP2C9 enzyme have lower dose requirements for glycemic control when compared to patients homozygous for wild-type alleles. Another important factor that potentially affects response to sulfonylureas is inhibition of the CYP2C9 enzyme due to concomitant administration of other drugs. Drug–gene interactions are the subject of clinical trial NCT01061112, in which investigators will assess the extent of CYP2C9 inhibition in tolbutamide-treated individuals with the *CYP2C9*1/*1*, *CYP2C9*1/*3*, and *CYP2C9*3/*3* genotypes by fluconazole.

30.5.2.2 Pharmacodynamics: Sulfonylurea 1 Receptor—Inward Rectifier K^+ Channel Kir6.2

The K_{ATP} channel in pancreatic B cells is the pharmacologic target of sulfonylureas. It is made up of four subunits of the sulfonylurea 1 receptor (SUR1), encoded by the *ABCC8* gene, and four subunits of the inward rectifier K^+ channel Kir6.2, encoded by *KCNJ11*. SUR1 is believed to be the regulatory subunit of the K_{ATP} channel, whereas Kir6.2 is the inwardly rectifying ion channel forming the pore and confers ATP sensitivity to the channel complex [51]. Gene polymorphisms of both components of the K_{ATP} channel have been associated with response to sulfonylurea therapy (Table 30.3).

The most important polymorphisms in SUR1 and Kir6.2 are the *ABCC8 S1369A* (rs757110) and *KCNJ11 E23K* (rs5219) variants, respectively. These polymorphisms are in almost complete LD and are inherited in linkage in more than 99% of haplotypes [52]. *KCNJ11 23K* affects *in vitro* K_{ATP} channel properties, inducing spontaneous hyperactivity of pancreatic B cells via increased threshold ATP concentration for insulin release. However, in the presence of sulfonylureas, *23K* K_{ATP} channels have increased sensitivity compared to *23E* K_{ATP} channels. Nevertheless, molecular electrophysiological evidence shows that *ABCC8 S1369A*, rather than *KCNJ11 E23K*, is the functional variant in the *ABCC8-KCNJ11* region [53]. More recent data suggest that K_{ATP} channel sensitivity is likely dependent on administered sulfonylurea [54]. K_{ATP} channels containing the *K23/A1369* haplotype are significantly less sensitive to inhibition by tolbutamide, chlorpropamide, and glimepiride, whereas K_{ATP} channels containing the *K23/A1369* haplotype are significantly more sensitive to inhibition by gliclazide [54].

Several studies have shown the association of *ABCC8-KCNJ11* variants with sulfonylurea response. In gliclazide-treated T2DM patients, it was shown that carriage of the variant alleles *ABCC8 1369A* and *KCNJ11 23K* leads to prolonged closure of K_{ATP} channel and insulin secretion, and that carriers of these alleles are more sensitive to sulfonylureas and have more efficient blood glucose and a greater decrease in HbA1c [53]. The results of a study of 115 T2DM patients treated with gliclazide showed that carriers of *ABCC8 1369A* exhibited greater significant HbA1c decrease after treatment compared to noncarriers (−1.6% versus −0.76%; $p = 0.044$) [55].

The association of the *ABCC8 1369A* allele with favorable response to sulfonylureas was replicated in a pooled analysis of two independent cohorts of gliclazide-treated T2DM patients [56]. In this study, patients with *ABCC8 1369SA* and *1369AA* genotypes, compared to *ABCC8 1369SS* patients, had greater decreases in fasting plasma glucose (−2.8% and −7.7%, respectively), greater decreases in 2-h glucose after oral glucose tolerance testing (−10.8% and −11.9%, respectively), and a greater chance of responding to gliclazide therapy compared to *ABCC8 1369SS* patients (OR 1.4 and 2.2, respectively) [56]. A similar effect of *KCNJ11 23K* (linked to *ABCC8 1369S*) on sulfonylurea response was reported [57]. In 101 T2DM patients who failed to achieve glycemic goals while on metformin monotherapy and were administered sulfonylureas as an add-on treatment, carriers of the *KCNJ11 23K* allele had greater decreases in HbA1c compared with *EE* individuals after six months of treatment. Interestingly, in a recent study of 100 Egyptian T2DM patients, the *KCNJ11 23K* allele was associated not with a favorable response, but with secondary sulfonylurea failure, defined as HbA1c ≥ 8%, despite sulfonylurea therapy [58].

ABCC8 1369A and *KCNJ11 23K* alleles have been studied also in relation to incidence of severe or mild hypoglycemia induced by sulfonylureas. In three studies these variant alleles were not associated with this major adverse effect of sulfonylureas [59–61].

30.5.2.3 Disease-Associated Gene Polymorphisms

As has already been discussed, the *TCF7L2* gene SNP rs7903146 (*C>T*) is the strongest predictor of T2DM incidence in populations of different ancestry. Carriers of the *TCF7L2* rs7903146 *T* allele have decreased B cell function compared to noncarriers. Along with this association, research has focused on the pharmacogenomic impact of this genomic marker to sulfonylurea therapeutic response. Results of studies conducted so far support the hypothesis that T2DM patients with decreased B cell function due to *TCF7L2* rs7903146 *T* carriage are at increased risk of sulfonylurea treatment failure (Table 30.4). In 901 sulfonylurea-treated T2DM participants of the Go-DARTs study, 57% of the rs7903146*TT* homozygotes failed to reach the glycemic target of HbA1c <7% compared to 40% of wild-type individuals [62]. Similar results were reported in two recent studies. In 87 T2DM patients who were on sulfonylurea + metformin therapy for six months, carriers of rs7903146 *T* (genotypes *CT* + *TT*) saw significantly smaller reduction in HbA1c and fasting plasma glucose in response to six months of sulfonylurea treatment [63], while in 189 T2DM patients who were treated with sulfonylureas, the *TCF7L2* genotype was the only predictor of sulfonylurea treatment failure (OR 1.57) [64].

In an ongoing clinical trial (NCT 00738088), the underlying genetic cause of differential response to sulfonylureas (gliclazide) is being assessed by identifying participants with a certain genetic predisposition to diabetes due to changes in the *TCF7L2* gene, then looking at how well they respond to sulfonylurea medication (Table 30.5).

30.5.3 Nonsulfonylurea Insulin Secretagogues (Glinides)

Nonsulfonylurea insulin secretagogues (glinides) have properties similar to the pharmacokinetic and pharmacodynamic properties of sulfonylureas; however, they bind to a different site in the sulfonylurea receptor to promote insulin secretion. Those currently available are repaglinide and nateglinide. The typical reductions in HbA1c levels they cause range from 0.5% to 1.5% [11].

30.5.3.1 Pharmacokinetics

Cytochrome P450 2C8 (CYP2C8) and 2C9 (CYP2C9)

The P450 2C8 (CYP2C8) enzyme is the primary metabolizer of repaglinide, whereas the CYP3A4 enzyme has only a minor contribution to its metabolism [36]. Polymorphisms in the *CYP2C8* gene affect the enzymatic activity of the encoded CYP2C8 protein and are associated with altered repaglinide pharmacokinetics. To date, in the Human Cytochrome P450 (CYP) Allele Nomenclature Database, 14 allelic variants have been identified in *CYP2C8* (www.cypalleles.ki.se/cyp2c8.htm). Among these, *CYP2C8*3* has been extensively studied in association with repaglinide pharmacokinetics and response (Table 30.2). Nateglinide is structurally similar to repaglinide; however, it is primarily metabolized by CYP2C9 enzyme [36]. Like sulfonylureas, the *CYP2C9*2* and **3* variant alleles are expected to affect nateglinide pharmacokinetics (Table 30.2).

In pharmacogenetic studies assessing the effect of gene polymorphisms on repaglinide pharmacokinetics, *CYP2C8*1/*3* individuals had lower mean AUC of repaglinide (−45%) and lower peak plasma concentration (−39%) compared with *CYP2C8*1/*1* individuals [65]. The research group involved in these studies replicated this association in a another study of 56 healthy volunteers [66]. In the case of nateglinide, *CYP2C9*3* allele carriers had significantly reduced oral nateglinide clearance; *CYP2C9*2/*3* individuals had median total clearances of 5.8 L/h; *CYP2C9*3/*3* individuals, 4.1 L/h; and *CYP2C9*1/*1* individuals, 7.9 L/h [67].

Similar results were obtained in a randomized, open-label, two-period crossover study of 18 healthy volunteers administered a single 120 mg nateglinide dose alone or in combination with sulfinpyrazone, a potent CYP2C9 inhibitor [68]. Co-administration of nateglinide with sulfinpyrazone resulted in higher mean nateglinide AUC (+28%) [68]. More recently, in a study of 35 Chinese male healthy volunteers administered a single oral dose of 120 mg nateglinide, the *CYP2C9*3* allele was associated with nateglinide AUC and clearance; however, no significant association with changes in the blood glucose-lowering effect of nateglinide was found [69].

30.5.3.2 Organic Anion-Transporting Polypeptide 1B1 (OATP1B1, SLCO1B1)

The disposition of both repaglinide and nateglinide in the human body is highly dependent on the polymorphic organic anion-transporting polypeptide 1B1 (OATP1B1), encoded by the *SLCO1B1* gene [70]. More than 40 nonsynonymous variants have been identified in this gene. The two most studied are rs4149056*T>C* (*521T>C*, Val174Ala) and rs2306283*A>G* (*388A>G*, Asn130Asp) (Table 30.1). The combination of these two functional SNPs forms four distinct haplotypes: *SLCO1B1*1A* (*388A-521T*), which is the reference haplotype, **1B* (*388G-521T*), **5* (*388A-521C*), and **15* (*388G-521C*) [71]. *In vitro* studies have shown that the *SLCO1B1*5* and *SLCO1B1*15* haplotypes (carriers of the *521C* allele) are associated with reduced OATP1B1 transport activity in different cell lines, suggesting that the effect of *521T>C* dominates that of *388A>G*.

One study investigated the effect of *SLCO1B1* gene polymorphisms for association with response to both repaglinide and nateglinide (Table 30.3). In the case of repaglinide, after analyzing the effect of *CYP2C8*3* on drug pharmacokinetics [65], investigators also assessed the effect of *SLCO1B1 −11187G>A*, *521T>C*, and *ABCB1 2677G>T/A* on drug pharmacokinetics in 56 healthy volunteers also genotyped for *CYP2C8* [66]. In another study, they replicated the results of the association of *CYP2C8*3* with a lower mean AUC of repaglinide and additionally showed that *SLCO1B1 521CC* individuals had higher AUC than *SLCO1B1 521TC* and *521TT* individuals (107% and 188%, respectively). In the latter study, both the *SLCO1B1 521CC* genotype and the *CYP2C8*3* allele were found to be independent predictors of AUC and concentration of repaglinide [66].

In another dose-escalation repaglinide study involving 28 healthy individuals, the *SLCO1B1* genotype was significantly associated with the pharmacokinetics of repaglinide in all tested doses [72]. After ingestion of a 0.25, 0.5, 1, or 2 mg dose of repaglinide, its AUC was 82, 72, 56, or 108% larger in *SLCO1B1 521CC* individuals compared with *SLCO1B1 521TT* individuals, whereas the maximum concentration was 60−70% larger in *SLCO1B1 521CC* individuals compared with *SLCO1B1 521TT* individuals. This same study also found that the effect of the *SLCO1B1* genotype on repaglinide plasma concentration translated into a tendency toward greater lowering of blood glucose concentrations in *SLCO1B1 521CC* individuals compared with *SLCO1B1 521TT* individuals [72].

More recently, the same investigators showed that, in 32 *CYP2C8*1/*1* healthy volunteers receiving a single 0.5 mg dose of repaglinide, the *SLCO1B1 521CC* individuals had a higher AUC than *SLCO1B1 521TC* and *521TT* individuals (59% and 72%, respectively) and a higher maximum concentration (27% and 62%, respectively) [73]. Similarly, increased repaglinide concentration was translated into a tendency in *SLCO1B1 521CC* individuals toward a greater blood glucose-lowering effect [73]. The *SLCO1B1 521T>C* polymorphism attenuated repaglinide clearance in 22 healthy male Chinese individuals; however, no significant differences in this effect were observed [74]. Overall, it appears that *CYP2C8* and *SLCO1B1* polymorphisms affect repaglinide pharmacokinetics and response.

In the case of nateglinide, in a study of 17 healthy volunteers with different *SLCO1B1* genotypes (11 individuals *521TT*, 4 individuals *521TC*, and 2 individuals *521CC*) administered an oral dose of 90 mg nateglinide, both maximum concentration and AUC were shown to be higher in *SLCO1B1 521CC* individuals (76% and 108%, respectively) and in *SLCO1B1 521TC* individuals (83% and 82%, respectively), compared with *SLCO1B1 521TT* individuals [75]. However, this finding was not replicated in another study [73]. In a more recent study involving 35 Chinese male healthy volunteers administered a single oral dose of 120 mg nateglinide, the association of not only *CYP2C9*3,* but also of *SLCO1B1 521 T>C* with AUC and clearance was demonstrated [69].

30.5.3.3 Pharmacodynamics: Sulfonylurea 1 Receptor—Inward Rectifier K$^+$ Channel Kir6.2 (KCNJ11)

Because both nonsulfonylurea insulin secretagogues, repaglinide and nateglinide, stimulate insulin secretion through bonding to the sulfonylurea receptor (SUR1, *ABCC8*) in B pancreatic cells, *ABCC8 S1370A* and inward rectifier K$^+$ channel Kir6.2 (*KCNJ11*) *E23K* polymorphisms have become attractive candidates that may affect response to these drugs. In a study of 100 Chinese patients newly diagnosed with T2DM and treated with repaglinide for 24 weeks, it was found that *KCNJ11 23K* is associated with better therapeutic effect (greater HbA1c decrease in *KCNJ11 23EK* and *23KK* individuals); the *ABCC8* exon/intron *16-3CC* homozygous genotype was associated with better insulin sensitivity [76]. Of 40 T2DM patients randomly selected to undergo an 8-week repaglinide treatment regimen, carriers of the *23E* allele had higher HbA1c levels compared to those of *23KK* patients [77]. Results on the association of the *ABCC8/KCNJ11* gene polymorphisms with nateglinide response are lacking. One possible explanation is that the inhibitory profile of the drug in K_{ATP} channels is similar irrespective of the variant haplotype [54].

Although genetic variants affecting repaglinide and nateglinide pharmacokinetics and pharmacodynamics are associated with drug response, we have not been able to find any ongoing or completed clinical trial assessing the impact of genetic polymorphism(s) on the therapeutic response of diabetic patients to these drugs.

30.5.4 Thiazolidinediones

Thiazolidinediones (or glitazones) are peroxisome-proliferator-activated receptor γ (PPARγ) agonists. PPARγ regulates fatty acid storage and glucose metabolism through modulation of various genes. The genes activated by it stimulate lipid uptake and adipogenesis by fat cells. Thiazolidinediones are the strongest ligands of PPARγ and are insulin sensitizers because they increase the sensitivity of muscles, fat, and liver to exogenous insulin. They lower HbA1c by 0.5−1.4% and appear to have a more durable effect on glycemic control compared to other oral antidiabetic drugs [11]. Originally, the clinically used thiazolidinediones comprised three drugs—namely, pioglitazone, rosiglitazone, and troglitazone. However, troglitazone was withdrawn from the market over a decade ago because of an idiosyncratic reaction leading to drug-induced hepatitis; more recently, rosiglitazone and its combination product with metformin were also officially withdrawn both by FDA and the European Medicines Agency (EMA) because of an apparent increase in cardiovascular risk. Pioglitazone is thus the only thiazolidinedione currently available in the market. However, pharmacogenomic studies have been conducted for all three thiazolidinediones and are included and discussed in the following sections (see also Tables 30.2–30.4, which summarize the pharmacogenomic studies on thiazolidinedione disposition and therapeutic response).

30.5.4.1 Pharmacokinetics: Cytochrome P450 2C8 (CYP2C8)

Thiazolidinediones are extensively metabolized by CYP450 enzymes. Pioglitazone is mainly metabolized by CYP2C8, whereas CYP3A4/5 and CYP1A1 contribute to a lesser extent [78]. Prior to pioglitazone's withdrawal, there had been only one study on the association of *CYP2C8* gene polymorphisms with pioglitazone pharmacokinetics. In this study, 16 healthy volunteers received the CYP2C8 strong inhibitor trimethoprim [79]. All received a single oral dose of 15 mg of pioglitazone while being treated either with placebo or with 160 mg of trimethoprim twice daily. Carriage of *CYP2C8*3* in the placebo arm was shown to lead to lower pioglitazone mean plasma concentrations, shorter pioglitazone half-life (higher metabolic ratio), and lower AUC. The AUC was 34% lower in *CYP2C8*3/*3* individuals and 26% lower in *CYP2C8*1/*3* individuals, compared to *CYP2C8*1/*1* individuals [79]. The association of trimethoprim with increased plasma concentrations of pioglitazone was also shown, although the *CYP2C8* genotype was not associated with it [79]. In 2013, three studies dealing with the pioglitazone–*CYP2C8* association were published. In 31 healthy volunteers who received a single 15 mg dose of pioglitazone followed by a 48-h sampling period, it was shown that *CYP2C8*3* had a significant effect on central compartment clearance [80].

In another two-armed study, 30 healthy Caucasian volunteers received either a single 15 mg dose of pioglitazone or 600 mg of the potent CYP2C8 inhibitor gemfibrozil twice daily for 4 days; they also received a single 15 mg dose of pioglitazone on the morning of day 3. It was shown that, when pioglitazone was administered alone, the mean AUC in *CYP2C8*3* carriers was 29.7% lower compared with that in *CYP2C8*1/*1* carriers [81]. When pioglitazone was coadministered with gemfibrozil, *CYP2C8*3* carriers had a 5.2-fold mean increase in AUC compared with a 3.3-fold mean AUC increase in *CYP2C8*1/*1* carriers. A similar effect was shown for the *CYP2C8*2* allele. A study of 17 healthy African-American volunteers (9 genotyped *CYP2C8*1/*1*, 7 genotyped *CYP2C8*1/*2*, and 1 genotyped *CYP2C8*2/*2*) demonstrated that the keto-pioglitazone metabolite–AUC ratio was significantly lower in *CYP2C8*2* carriers than in *CYP2C8*1/*1* carriers [82].

Rosiglitazone is mainly metabolized by CYP2C8, with a minor contribution from CYP2C9 [83]. In a seminal randomized, crossover study investigating the effect of trimethoprim, a selective CYP2C8 inhibitor, on the CYP2C8-mediated metabolism of rosiglitazone, eight healthy individuals received a single 8 mg dose of the drug in the presence or absence of trimethoprim [84]. The study showed that trimethoprim inhibited rosiglitazone metabolism both *in vitro* and *in vivo*. However, in contrast with the previously described pioglitazone, the *CYP2C8* genotype (**2* or **3* allele) did not appear to affect rosiglitazone metabolism, although there was a trend toward a higher AUC in *CYP2C8*1/*1* carriers compared with carriers of *CYP2C8*2* or **3* [84]. Similarly, in a two-phase, open-label, crossover trial involving 23 healthy volunteers, aimed at determining the effect of *CYP2C8* and of fluvoxamine (a CYP2C8 inhibitor) on the pharmacokinetics of rosiglitazone, even though fluvoxamine caused a significant increase in AUC, no effect of *CYP2C8* in rosiglitazone metabolism was found [85].

In a larger clinical study of 31 healthy volunteers genotyped for *CYP2C8* and receiving a single dose of rosiglitazone (8 mg), carriers of *CYP2C8*3* showed higher *in vivo* metabolic capacity compared to *CYP2C8*1/*1* carriers. The study population consisted of 14 *CYP2C8*1/*1* individuals, 13 *CYP2C8*1/*3* individuals, and 4 *CYP2C8*3/*3* individuals. Researchers found that mean total rosiglitazone and desmethylrosiglitazone clearance was higher in *CYP2C8*3* carriers (*CYP2C8*1/*3* and *CYP2C8*3/*3*), whereas, accordingly, rosiglitazone elimination half-lives were lower in *CYP2C8*1/*3* and *CYP2C8*3/*3*, compared with *CYP2C8*1/*1* individuals [86]. However, no differences were observed between the glucose-lowering effect of rosiglitazone and the *CYP2C8* genotype [86]. In a subsequent study of 26 healthy Caucasian volunteers, subjects with the *CYP2C8*1/*3* genotype had significantly lower rosiglitazone AUC and significantly higher rosiglitazone oral clearance compared with *CYP2C8*1/*1* individuals [87].

In a more recent study of 187 rosiglitazone-treated patient participants in the South Danish Diabetes Study, carriers of the *CYP2C8*3* allele had lower plasma levels of rosiglitazone and reduced therapeutic response [88]. In the same study, *CYP2C8*3* carriers also had a lower risk of developing edema during treatment with rosiglitazone. Apart from the *CYP2C8*2* and **3* alleles, a novel allele, *CYP2C8*11*, was also found to be associated with rosiglitazone plasma concentration in 9 healthy volunteers of Asian origin treated with a single 4 mg dose of rosiglitazone [89].

30.5.4.2 Organic Anion-Transporting Polypeptide 1B1 (OATP1B1, SLCO1B1)

Pioglitazone and rosiglitazone are possible substrates of organic anion-transporting polypeptide 1B1, or OATP1B1 [90]. In a study of 32 healthy volunteers receiving a single 15 mg dose of pioglitazone, lack of association between the *SLCO1B1 521T>C* polymorphism and pioglitazone pharmacokinetics was the main finding [91]. In the same study, it was shown that, similarly to pioglitazone, no association was found between *SLCO1B1 521T>C* and rosiglitazone pharmacokinetics volunteers receiving a single 4 mg dose of rosiglitazone [91]. This is in agreement with another study showing a lack of association between *SLCO1B1 521T>C*, *−11187G>A*, *−10499A>C*, and *388A>G* and rosiglitazone pharmacokinetics in 26 healthy Caucasian volunteers with known *CYP2C8* genotype [87].

30.5.4.3 Pharmacodynamics: Peroxisome-Proliferator-Activated Receptor γ

Since thiazolidinediones are ligands of PPARγ, *PPARγ* gene polymorphisms are attractive candidates for uncovering novel pharmacogenetic markers related to interindividual variability in glitazone drug response. *PPARγ* is a transcription factor that, upon activation, heterodimerizes with the retinoid X receptor and binds to specific PPAR-responsive elements of DNA to promote transcription of numerous target genes [92]. A number of genetic variants in the *PPARγ* gene have been identified; the most promising association between *PPARγ* gene polymorphisms and thiazolidinediones response is with *PPARγ Pro12Ala*. The *12Ale* allele leads to reduced binding to the *PPARγ*-responsive DNA elements and indicates reduced transcriptional activity of *PPARγ* [92].

In the case of pioglitazone, Ramirez-Salazar et al. assessed the influence of *PPARγ Pro12Ala* on metabolic and hormonal response to pioglitazone treatment in 83 obese postmenopausal women treated with 15 mg/day of pioglitazone for 15 days [93]. The researchers were able to show that the *PPARγ 12ProAla* genotype is associated with a greater decrease in serum glucose levels compared with the *12ProPro* genotype (−15 mg/dL versus −7 mg/dL, p = 0.003). Therefore, it was concluded that obese women carrying the *12Ala* allele may benefit more from pioglitazone treatment [93].

The beneficial effect of the *PPARγ 12Ala* allele was also shown in a subsequent study of 250 T2DM Chinese patients treated with 30 mg of pioglitazone for 24 weeks. The *PPARγ 12Ala* allele (*12ProAla* and *12AlaAla* genotypes) was associated with improved therapeutic response to pioglitazone [94]. Carriers of *PPARγ 12Ala* were more likely to be pioglitazone responders than noncarriers. In the same study, carriers of the *12Ala* allele had significantly greater decrease in HbA1c and fasting glucose plasma levels compared with noncarriers after 24 weeks of pioglitazone treatment [94].

In another study of 101 Iranian T2DM patients treated with pioglitazone 15 mg/day for 12 weeks, carriers of the *12Ala* variant appeared to have a better therapeutic response than *12ProPro* individuals, although the difference between groups did not reach statistical significance [95]. In a more recent study of 67 Chinese T2DM patients receiving pioglitazone 30 mg/day for 3 months, carriers of the *12Ala* allele showed significantly higher differential values of postprandial plasma glucose and serum triglyceride compared with *12ProPro* carriers [96]. However, in another study of 30 T2DM patients of Asian Indian origin receiving pioglitazone monotherapy, the *PPARγ Pro12Ala* polymorphism was not associated with glycemic target or primary pioglitazone therapy failure [97].

Only one study is available for the association of *PPARγ Pro12Ala* polymorphism with rosiglitazone response. As with pioglitazone, in a study of 198 T2DM patients treated with 4 mg/day of rosiglitazone for 12 weeks, the *12Ala* allele was associated with a greater decrease both in fasting plasma glucose levels and in HbA1c levels [98].

Before troglitazone was withdrawn, two randomized clinical trials—Troglitazone in the Prevention of Diabetes (TRIPOD) and DPP—assessed its effect on the prevention of diabetes. Participants of both trials were genotyped for *PPARγ* variants. Interestingly, the results from both cohorts revealed the *PPARγ Pro12Ala* polymorphism to be a major determinant for troglitazone therapeutic response [99,100]. Consistently, *12Ala* carriers appeared more likely to benefit from troglitazone therapy [99,100]. This finding strengthens the hypothesis that *PPARγ Pro12Ala* and other genetic variants of the *PPARγ* gene are associated with therapeutic response to thiazolidinediones, and thus might be useful genetic markers for identifying individuals who can benefit from therapy with the currently available thiazolidinedione, pioglitazone; however, it cannot resolve safety concerns about troglitazone's association with hepatotoxicity, or that of rosiglitazone with cardiovascular events.

A clinical trial aiming at determining genetic predictors of response to pioglitazone (NCT01135394) is currently recruiting patients (Table 30.5). T2DM patients will receive 30 mg pioglitazone once daily for 4 weeks and 45 mg once daily for an additional 8 weeks. Among the trial's outcome measures are the identification of genes correlated to pioglitazone response and polymorphisms and haplotypes in candidate genes that influence this response (Table 30.5). Given the heavy risk profile of thiazolidinediones, the availability of pharmacogenetic markers that could be used to stratify patients and predict response to thiazolidinediones, especially pioglitazone, which is the only thiazolidinedione currently available, is essential to improve and guide antidiabetic therapeutics with this drug.

30.6 STRATIFIED ORAL ANTIDIABETIC MEDICINE: PHARMACOGENOMICS APPLICATION

According to clinical guidelines, lifestyle modifications followed by pharmacotherapy with metformin are appropriate initial treatments for most individuals diagnosed with T2DM. When glycemic goals are not achieved, addition of a further antidiabetic agent, including a sulfonylurea, a meglitinide, or pioglitazone is often the second step. The use of intensive glucose-lowering therapy needs to be tailored to individual patient needs and risks. Maintaining effective glycemic control is linked to a diminished risk of microvascular complications. Currently, the choice of antidiabetic regimen relies on demographic and clinical characteristics of T2DM patients, whereas HbA1c monitoring directs the required therapy modifications.

Several of the genetic markers described in this chapter show an association with therapeutic response and/or adverse effects of oral antidiabetic drugs, and they point toward the utility of pharmacogenomics in personalizing and optimizing oral antidiabetic therapy. The pharmacogenomics of oral antidiabetic medications has the potential to improve clinical practice in terms of choosing the most appropriate oral hypoglycemic agent at the optimal dose based on individual genetic profiles. Genotyping of a panel of gene polymorphisms associated with response to metformin, sulfonylureas, nonsulfonylurea secretagogues, and pioglitazone can lead to instant diabetes control and reduced incidence of adverse events. Use of oral antidiabetic medication pharmacogenomics at the initiation of pharmacotherapy can also help in long-term reduction of micro- and macrovascular complications and in cardiovascular risk reduction. Ultimately, a pharmacogenomic-guided treatment approach will not only increase overall efficacy and the likelihood of a patient adherence, but may also improve patient well-being and quality of life.

Currently, pharmacogenomic information is not incorporated in antidiabetic drug inserts. Two main limitations of oral antidiabetic pharmacogenomics in routine clinical practice are the lack of large prospective clinical trials showing the benefits of pharmacogenomics at therapy initiation and the availability of many different pharmacologic compounds that enable clinicians to switch among different therapies. Although achieving and maintaining glycemic goals in T2DM patients is crucial for disease control, it should be acknowledged that no initial therapy failure has the devastating results seen in other life-threatening diseases such as cancer. Physicians appear content with the current empirical approach of starting and adjusting antidiabetic therapy as necessary, using established nongenetic indicators. However, it may become increasingly difficult to continue with this approach as more evidence leads to the era of personalized medicine.

30.7 CONCLUSIONS: FUTURE PERSPECTIVES

It is anticipated that the results of ongoing clinical trials assessing the pharmacogenomic dimension of oral antidiabetic drugs, as summarized in Table 30.5, will strengthen currently available evidence on the usefulness of oral antidiabetic pharmacogenomics and help persuade physicians to start using this tool. Of course, several more trials will be necessary to consolidate any favorable findings of completed trials. The ultimate trials will be those testing algorithms that incorporate pharmacogenetic information provided by patient genotype, along with other nongenomic stratification parameters, and that produce specific dose instructions that give tangible clinical benefit to patients. These trials should lead to the incorporation of relevant pharmacogenetic information in antidiabetic drug inserts and in therapeutic guidelines of medical associations. Ultimately, to get to the clinic, the single most important step necessary, in our opinion, will be the education of clinicians in using these tools in their everyday routine practice. For that, we still have a long, uphill way to go.

ABBREVIATIONS

AACE	American Association of Clinical Endocrinologists
ACE	American College of Endocrinology
ADA	American Diabetes Association
ATM	ataxia telangiectasia mutated
AUC	area under the concentration–time curve
CARDS	Collaborative Atorvastatin Diabetes Study
CYP450	cytochrome P450
DCS	diabetes care system
DPP	Diabetes Prevention Program
EASD	European Association for the Study of Diabetes
EM	extensive metabolizer
EMA	European Medicines Agency
FDA	U.S. Food and Drug Administration
Go-DARTs	Genetics of Diabetes Audit and Research Tayside
GWAS	genome-wide association study
HbA1c	glycated hemoglobin
IM	intermediate metabolizer
K_{ATP} channel	ATP-dependent potassium channel
LD	linkage disequilibrium
MAF	minor allele frequency
MATE1	multidrug and toxin extrusion 1
OATP1B1	organic anion-transporting polypeptide 1B1
OCT	organic cation transporter
OR	odds ratio
PM	poor metabolizer
PPARγ	peroxisome proliferator-activated receptor γ
SNP	single-nucleotide polymorphism
SUR1	sulfonylurea 1 receptor
T2DM	type 2 diabetes mellitus
TCF7L2	transcription factor 7-like 2
TRIPOD	Troglitazone in the Prevention of Diabetes
UKPDS	UK Prospective Diabetes Study

REFERENCES

[1] Fowler MJ. Microvascular and macrovascular complications of diabetes. Clin Diab 2008;26(2):77–82.

[2] Imam K. Clinical features, diagnostic criteria and pathogenesis of diabetes mellitus. Adv Exp Med Biol 2012;771:340–55.

[3] Ahlqvist E, Ahluwalia TS, Groop L. Genetics of type 2 diabetes. Clin Chem 2011;57(2):241–54.

[4] Marchetti P, Syed F, Suleiman M, Bugliani M, Marselli L. From genotype to human beta cell phenotype and beyond. Islets 2012;4(5):323–32.

[5] Iordanidou M, Tavridou A, Petridis I, Arvanitidis KI, Christakidis D, Vargemezis V, et al. The serotonin transporter promoter polymorphism (5-HTTLPR) is associated with type 2 diabetes. Clin Chim Acta 2010;411(3–4):167–71.

[6] Hirschhorn JN, Gajdos ZK. Genome-wide association studies: results from the first few years and potential implications for clinical medicine. Annu Rev Med 2011;62:11–24.

[7] Hansson O, Zhou Y, Renstrom E, Osmark P. Molecular function of TCF7L2: consequences of TCF7L2 splicing for molecular function and risk for type 2 diabetes. Curr Diab Rep 2010;10(6):444–51.

[8] Franklin CS, Aulchenko YS, Huffman JE, Vitart V, Hayward C, Polasek O, et al. The TCF7L2 diabetes risk variant is associated with HbA(C) levels: a genome-wide association meta-analysis. Ann Hum Genet 2010;74(6):471–8.

[9] Csoka AB, Szyf M. Epigenetic side-effects of common pharmaceuticals: a potential new field in medicine and pharmacology. Med Hypotheses 2009;73(5):770–80.

[10] Manolopoulos VG, Ragia G, Tavridou A. Pharmacogenomics of oral antidiabetic medications: current data and pharmacoepigenomic perspective. Pharmacogenomics 2011;12(8):1161–91.

[11] Nathan DM, Buse JB, Davidson MB, Ferrannini E, Holman RR, Sherwin R, et al. Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy: a consensus statement of the American diabetes association and the European association for the study of diabetes. Diabetes Care 2009;32(1):193–203.

[12] Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, et al. Management of hyperglycemia in type 2 diabetes: a

patient-centered approach: position statement of the American diabetes association (ADA) and the European association for the study of diabetes (EASD). Diabetes Care 2012;35(6):1364–79.

[13] Manolopoulos VG. Pharmacogenomics and adverse drug reactions in diagnostic and clinical practice. Clin Chem Lab Med 2007;45(7):801–14.

[14] Gong L, Goswami S, Giacomini KM, Altman RB, Klein TE. Metformin pathways: pharmacokinetics and pharmacodynamics. Pharmacogenet Genomics 2012;22(11):820–7.

[15] Shikata E, Yamamoto R, Takane H, Shigemasa C, Ikeda T, Otsubo K, et al. Human organic cation transporter (OCT1 and OCT2) gene polymorphisms and therapeutic effects of metformin. J Hum Genet 2007;52(2):117–22.

[16] Kerb R, Brinkmann U, Chatskaia N, Gorbunov D, Gorboulev V, Mornhinweg E, et al. Identification of genetic variations of the human organic cation transporter hOCT1 and their functional consequences. Pharmacogenetics 2002;12(8):591–5.

[17] Shu Y, Sheardown SA, Brown C, Owen RP, Zhang S, Castro RA, et al. Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. J Clin Invest 2007;117(5):1422–31.

[18] Shu Y, Brown C, Castro RA, Shi RJ, Lin ET, Owen RP, et al. Effect of genetic variation in the organic cation transporter 1, OCT1, on metformin pharmacokinetics. Clin Pharmacol Ther 2008;83(2):273–80.

[19] Tzvetkov MV, Vormfelde SV, Balen D, Meineke I, Schmidt T, Sehrt D, et al. The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3 on the renal clearance of metformin. Clin Pharmacol Ther 2009;86(3):299–306.

[20] Becker ML, Visser LE, van Schaik RH, Hofman A, Uitterlinden AG, Stricker BH. Genetic variation in the organic cation transporter 1 is associated with metformin response in patients with diabetes mellitus. Pharmacogenomics J 2009;9(4):242–7.

[21] Christensen MM, Brasch-Andersen C, Green H, Nielsen F, Damkier P, Beck-Nielsen H, et al. The pharmacogenetics of metformin and its impact on plasma metformin steady-state levels and glycosylated hemoglobin A1c. Pharmacogenet Genomics 2011;21(12):837–50.

[22] Yoon H, Cho HY, Yoo HD, Kim SM, Lee YB. Influences of organic cation transporter polymorphisms on the population pharmacokinetics of metformin in healthy subjects. Aaps J 2013;15(2):571–80.

[23] Tarasova L, Kalnina I, Geldnere K, Bumbure A, Ritenberga R, Nikitina-Zake L, et al. Association of genetic variation in the organic cation transporters OCT1, OCT2 and multidrug and toxin extrusion 1 transporter protein genes with the gastrointestinal side effects and lower BMI in metformin-treated type 2 diabetes patients. Pharmacogenet Genomics 2012;22(9):659–66.

[24] Tkac I, Klimcakova L, Javorsky M, Fabianova M, Schroner Z, Hermanova H, et al. Pharmacogenomic association between a variant in SLC47A1 gene and therapeutic response to metformin in type 2 diabetes. Diabetes Obes Metab 2013;15(2):189–91.

[25] Song IS, Shin HJ, Shim EJ, Jung IS, Kim WY, Shon JH, et al. Genetic variants of the organic cation transporter 2 influence the disposition of metformin. Clin Pharmacol Ther 2008;84(5):559–62.

[26] Wang ZJ, Yin OQ, Tomlinson B, Chow MS. OCT2 polymorphisms and *in-vivo* renal functional consequence: studies with metformin and cimetidine. Pharmacogenet Genomics 2008;18(7):637–45.

[27] Chen Y, Li S, Brown C, Cheatham S, Castro RA, Leabman MK, et al. Effect of genetic variation in the organic cation transporter 2 on the renal elimination of metformin. Pharmacogenet Genomics 2009;19(7):497–504.

[28] Damme K, Nies AT, Schaeffeler E, Schwab M. Mammalian MATE (SLC47A) transport proteins: impact on efflux of endogenous substrates and xenobiotics. Drug Metab Rev 2011;43(4):499–523.

[29] Becker ML, Visser LE, van Schaik RH, Hofman A, Uitterlinden AG, Stricker BH. Genetic variation in the multidrug and toxin extrusion 1 transporter protein influences the glucose-lowering effect of metformin in patients with diabetes: a preliminary study. Diabetes 2009;58(3):745–9.

[30] Becker ML, Visser LE, van Schaik RH, Hofman A, Uitterlinden AG, Stricker BH. Interaction between polymorphisms in the OCT1 and MATE1 transporter and metformin response. Pharmacogenet Genomics 2010;20(1):38–44.

[31] Jablonski KA, McAteer JB, de Bakker PI, Franks PW, Pollin TI, Hanson RL, et al. Common variants in 40 genes assessed for diabetes incidence and response to metformin and lifestyle intervention in the diabetes prevention program. Diabetes 2010;59(10):2672–81.

[32] Zhou K, Bellenguez C, Spencer CC, Bennett AJ, Coleman RL, Tavendale R, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. Nat Genet 2011;43(2):117–20.

[33] van Leeuwen N, Nijpels G, Becker ML, Deshmukh H, Zhou K, Stricker BH, et al. A gene variant near ATM is significantly associated with metformin treatment response in type 2 diabetes: a replication and meta-analysis of five cohorts. Diabetologia 2012;55(7):1971–7.

[34] Florez JC, Jablonski KA, Taylor A, Mather K, Horton E, White NH, et al. The C allele of ATM rs11212617 does not associate with metformin response in the diabetes prevention program. Diabetes Care 2012;35(9):1864–7.

[35] Rodbard HW, Jellinger PS, Davidson JA, Einhorn D, Garber AJ, Grunberger G, et al. Statement by an American association of clinical endocrinologists/American college of endocrinology consensus panel on type 2 diabetes mellitus: an algorithm for glycemic control. Endocr Pract 2009;15(6):540–59.

[36] Kirchheiner J, Roots I, Goldammer M, Rosenkranz B, Brockmoller J. Effect of genetic polymorphisms in cytochrome p450 (CYP) 2C9 and CYP2C8 on the pharmacokinetics of oral antidiabetic drugs: clinical relevance. Clin Pharmacokinet 2005;44(12):1209–25.

[37] Niemi M, Cascorbi I, Timm R, Kroemer HK, Neuvonen PJ, Kivisto KT. Glyburide and glimepiride pharmacokinetics in subjects with different CYP2C9 genotypes. Clin Pharmacol Ther 2002;72(3):326–32.

[38] Kirchheiner J, Bauer S, Meineke I, Rohde W, Prang V, Meisel C, et al. Impact of CYP2C9 and CYP2C19 polymorphisms on tolbutamide kinetics and the insulin and glucose response in healthy volunteers. Pharmacogenetics 2002;12(2):101–9.

[39] Kirchheiner J, Brockmoller J, Meineke I, Bauer S, Rohde W, Meisel C, et al. Impact of CYP2C9 amino acid polymorphisms on glyburide kinetics and on the insulin and glucose response in healthy volunteers. Clin Pharmacol Ther 2002;71(4):286–96.

[40] Lee CR, Pieper JA, Hinderliter AL, Blaisdell JA, Goldstein JA. Evaluation of cytochrome P4502C9 metabolic activity with tolbutamide in CYP2C91 heterozygotes. Clin Pharmacol Ther 2002;72(5):562–71.

[41] Tan B, Zhang YF, Chen XY, Zhao XH, Li GX, Zhong DF. The effects of CYP2C9 and CYP2C19 genetic polymorphisms on the pharmacokinetics and pharmacodynamics of glipizide in Chinese subjects. Eur J Clin Pharmacol 2010;66(2):145–51.

[42] Becker ML, Visser LE, Trienekens PH, Hofman A, van Schaik RH, Stricker BH. Cytochrome P450 2C9 *2 and *3 polymorphisms and the dose and effect of sulfonylurea in type II diabetes mellitus. Clin Pharmacol Ther 2008;83(2):288–92.

[43] Zhou K, Donnelly L, Burch L, Tavendale R, Doney AS, Leese G, et al. Loss-of-function CYP2C9 variants improve therapeutic response to sulfonylureas in type 2 diabetes: a Go-DARTS study. Clin Pharmacol Ther 2010;87(1):52–6.

[44] Swen JJ, Wessels JA, Krabben A, Assendelft WJ, Guchelaar HJ. Effect of CYP2C9 polymorphisms on prescribed dose and time-to-stable dose of sulfonylureas in primary care patients with Type 2 diabetes mellitus. Pharmacogenomics 2010;11(11):1517–23.

[45] Surendiran A, Pradhan SC, Agrawal A, Subrahmanyam DK, Rajan S, Anichavezhi D, et al. Influence of CYP2C9 gene polymorphisms on response to glibenclamide in type 2 diabetes mellitus patients. Eur J Clin Pharmacol 2011;67(8):797–801.

[46] Chao J, Nau DP, Aikens JE. Patient-reported perceptions of side effects of antihyperglycemic medication and adherence to medication regimens in persons with diabetes mellitus. Clin Ther 2007;29(1):177–80.

[47] Holstein A, Plaschke A, Ptak M, Egberts EH, El-Din J, Brockmoller J, et al. Association between CYP2C9 slow metabolizer genotypes and severe hypoglycaemia on medication with sulphonylurea hypoglycaemic agents. Br J Clin Pharmacol 2005;60(1):103–6.

[48] Ragia G, Petridis I, Tavridou A, Christakidis D, Manolopoulos VG. Presence of CYP2C9*3 allele increases risk for hypoglycemia in Type 2 diabetic patients treated with sulfonylureas. Pharmacogenomics 2009;10(11):1781–7.

[49] Gokalp O, Gunes A, Cam H, Cure E, Aydin O, Tamer MN, et al. Mild hypoglycaemic attacks induced by sulphonylureas related to CYP2C9, CYP2C19 and CYP2C8 polymorphisms in routine clinical setting. Eur J Clin Pharmacol 2011;67(12):1223–9.

[50] Holstein A, Hahn M, Patzer O, Seeringer A, Kovacs P, Stingl J. Impact of clinical factors and CYP2C9 variants for the risk of severe sulfonylurea-induced hypoglycemia. Eur J Clin Pharmacol 2011;67(5):471–6.

[51] Miki T, Nagashima K, Seino S. The structure and function of the ATP-sensitive K+ channel in insulin-secreting pancreatic beta-cells. J Mol Endocrinol 1999;22(2):113–23.

[52] Florez JC, Burtt N, de Bakker PI, Almgren P, Tuomi T, Holmkvist J, et al. Haplotype structure and genotype-phenotype correlations of the sulfonylurea receptor and the islet ATP-sensitive potassium channel gene region. Diabetes 2004;53(5):1360–8.

[53] Hamming KS, Soliman D, Matemisz LC, Niazi O, Lang Y, Gloyn AL, et al. Coexpression of the type 2 diabetes susceptibility gene variants KCNJ11 E23K and ABCC8 S1369A alter the ATP and sulfonylurea sensitivities of the ATP-sensitive K(+) channel. Diabetes 2009;58(10):2419–24.

[54] Lang VY, Fatehi M, Light PE. Pharmacogenomic analysis of ATP-sensitive potassium channels coexpressing the common type 2 diabetes risk variants E23K and S1369A. Pharmacogenet Genomics 2012;22(3):206–14.

[55] Zhang H, Liu X, Kuang H, Yi R, Xing H. Association of sulfonylurea receptor 1 genotype with therapeutic response to gliclazide in type 2 diabetes. Diabetes Res Clin Pract 2007;77(1):58–61.

[56] Feng Y, Mao G, Ren X, Xing H, Tang G, Li Q, et al. Ser1369Ala variant in sulfonylurea receptor gene ABCC8 is associated with antidiabetic efficacy of gliclazide in Chinese type 2 diabetic patients. Diabetes Care 2008;31(10):1939–44.

[57] Javorsky M, Klimcakova L, Schroner Z, Zidzik J, Babjakova E, Fabianova M, et al. KCNJ11 gene E23K variant and therapeutic response to sulfonylureas. Eur J Intern Med 2012;23(3):245–9.

[58] El-Sisi AE, Hegazy SK, Metwally SS, Wafa AM, Dawood NA. Effect of genetic polymorphisms on the development of secondary failure to sulfonylurea in Egyptian patients with type 2 diabetes. Ther Adv Endocrinol Metab 2012;2(4):155–64.

[59] Ragia G, Tavridou A, Petridis I, Manolopoulos VG. Association of KCNJ11 E23K gene polymorphism with hypoglycemia in sulfonylurea-treated type 2 diabetic patients. Diabetes Res Clin Pract 2012;98(1):119–24.

[60] Sato R, Watanabe H, Genma R, Takeuchi M, Maekawa M, Nakamura H. ABCC8 polymorphism (Ser1369Ala): influence on severe hypoglycemia due to sulfonylureas. Pharmacogenomics 2010;11(12):1743–50.

[61] Holstein A, Hahn M, Stumvoll M, Kovacs P. The E23K variant of KCNJ11 and the risk for severe sulfonylurea-induced hypoglycemia in patients with type 2 diabetes. Horm Metab Res 2009;41(5):387–90.

[62] Pearson ER, Donnelly LA, Kimber C, Whitley A, Doney AS, McCarthy MI, et al. Variation in TCF7L2 influences therapeutic response to sulfonylureas: a GoDARTs study. Diabetes 2007;56(8):2178–82.

[63] Schroner Z, Javorsky M, Tkacova R, Klimcakova L, Dobrikova M, Habalova V, et al. Effect of sulphonylurea treatment on glycaemic control is related to TCF7L2 genotype in patients with Type 2 diabetes. Diabetes Obes Metab 2011;13(1):89–91.

[64] Holstein A, Hahn M, Korner A, Stumvoll M, Kovacs P. TCF7L2 and therapeutic response to sulfonylureas in patients with type 2 diabetes. BMC Med Genet 2011;12:30.

[65] Niemi M, Leathart JB, Neuvonen M, Backman JT, Daly AK, Neuvonen PJ. Polymorphism in CYP2C8 is associated with reduced plasma concentrations of repaglinide. Clin Pharmacol Ther 2003;74(4):380–7.

[66] Niemi M, Backman JT, Kajosaari LI, Leathart JB, Neuvonen M, Daly AK, et al. Polymorphic organic anion transporting polypeptide 1B1 is a major determinant of repaglinide pharmacokinetics. Clin Pharmacol Ther 2005;77(6):468–78.

[67] Kirchheiner J, Meineke I, Muller G, Bauer S, Rohde W, Meisel C, et al. Influence of CYP2C9 and CYP2D6 polymorphisms on the pharmacokinetics of nateglinide in genotyped healthy volunteers. Clin Pharmacokinet 2004;43(4):267–78.

[68] Sabia H, Sunkara G, Ligueros-Saylan M, Wang Y, Smith H, McLeod J, et al. Effect of a selective CYP2C9 inhibitor on the pharmacokinetics of nateglinide in healthy subjects. Eur J Clin Pharmacol 2004;60(6):407–12.

[69] Cheng Y, Wang G, Zhang W, Fan L, Chen Y, Zhou HH. Effect of CYP2C9 and SLCO1B1 polymorphisms on the pharmacokinetics and pharmacodynamics of nateglinide in healthy Chinese male volunteers. Eur J Clin Pharmacol 2013;69(3):407–13.

[70] Kalliokoski A, Neuvonen PJ, Niemi M. SLCO1B1 polymorphism and oral antidiabetic drugs. Basic Clin Pharmacol Toxicol 2010;107(4):775–81.

[71] Tirona RG, Leake BF, Merino G, Kim RB. Polymorphisms in OATP-C: identification of multiple allelic variants associated with altered transport activity among European- and African-Americans. J Biol Chem 2001;276(38):35669–75.

[72] Kalliokoski A, Neuvonen M, Neuvonen PJ, Niemi M. The effect of SLCO1B1 polymorphism on repaglinide pharmacokinetics persists over a wide dose range. Br J Clin Pharmacol 2008;66(6):818–25.

[73] Kalliokoski A, Neuvonen M, Neuvonen PJ, Niemi M. Different effects of SLCO1B1 polymorphism on the pharmacokinetics and pharmacodynamics of repaglinide and nateglinide. J Clin Pharmacol 2008;48(3):311–21.

[74] He J, Qiu Z, Li N, Yu Y, Lu Y, Han D, et al. Effects of SLCO1B1 polymorphisms on the pharmacokinetics and pharmacodynamics of repaglinide in healthy Chinese volunteers. Eur J Clin Pharmacol 2011;67(7):701–7.

[75] Zhang W, He YJ, Han CT, Liu ZQ, Li Q, Fan L, et al. Effect of SLCO1B1 genetic polymorphism on the pharmacokinetics of nateglinide. Br J Clin Pharmacol 2006;62(5):567–72.

[76] He YY, Zhang R, Shao XY, Hu C, Wang CR, Lu JX, et al. Association of KCNJ11 and ABCC8 genetic polymorphisms with response to repaglinide in Chinese diabetic patients. Acta Pharmacol Sin 2008;29(8):983–9.

[77] Yu M, Xu XJ, Yin JY, Wu J, Chen X, Gong ZC, et al. KCNJ11 Lys23Glu and TCF7L2 rs290487(C/T) polymorphisms affect therapeutic efficacy of repaglinide in Chinese patients with type 2 diabetes. Clin Pharmacol Ther 2010;87(3):330–5.

[78] Jaakkola T, Laitila J, Neuvonen PJ, Backman JT. Pioglitazone is metabolised by CYP2C8 and CYP3A4 *in vitro*: potential for interactions with CYP2C8 inhibitors. Basic Clin Pharmacol Toxicol 2006;99(1):44–51.

[79] Tornio A, Niemi M, Neuvonen PJ, Backman JT. Trimethoprim and the CYP2C8*3 allele have opposite effects on the pharmacokinetics of pioglitazone. Drug Metab Dispos 2008;36(1):73–80.

[80] Kadam R, Bourne D, Kompella U, Aquilante C. Effect of cytochrome P450 2C8*3 on the population pharmacokinetics of pioglitazone in healthy Caucasian volunteers. Biol Pharm Bull 2013;36(2):245–51.

[81] Aquilante CL, Kosmiski LA, Bourne DW, Bushman LR, Daily EB, Hammond KP, et al. Impact of the CYP2C8 *3 polymorphism on the drug-drug interaction between gemfibrozil and pioglitazone. Br J Clin Pharmacol 2013;75(1):217–26.

[82] Aquilante CL, Wempe MF, Spencer SH, Kosmiski LA, Predhomme JA, Sidhom MS. Influence of CYP2C8*2 on the pharmacokinetics of pioglitazone in healthy African-American volunteers. Pharmacotherapy 2013;33(9):1000–7.

[83] Baldwin SJ, Clarke SE, Chenery RJ. Characterization of the cytochrome P450 enzymes involved in the *in vitro* metabolism of rosiglitazone. Br J Clin Pharmacol 1999;48(3):424–32.

[84] Hruska MW, Amico JA, Langaee TY, Ferrell RE, Fitzgerald SM, Frye RF. The effect of trimethoprim on CYP2C8 mediated rosiglitazone metabolism in human liver microsomes and healthy subjects. Br J Clin Pharmacol 2005;59(1):70–9.

[85] Pedersen RS, Damkier P, Brosen K. The effects of human CYP2C8 genotype and fluvoxamine on the pharmacokinetics of rosiglitazone in healthy subjects. Br J Clin Pharmacol 2006;62(6):682–9.

[86] Kirchheiner J, Thomas S, Bauer S, Tomalik-Scharte D, Hering U, Doroshyenko O, et al. Pharmacokinetics and pharmacodynamics of rosiglitazone in relation to CYP2C8 genotype. Clin Pharmacol Ther 2006;80(6):657–67.

[87] Aquilante CL, Bushman LR, Knutsen SD, Burt LE, Rome LC, Kosmiski LA. Influence of SLCO1B1 and CYP2C8 gene polymorphisms on rosiglitazone pharmacokinetics in healthy volunteers. Hum Genomics 2008;3(1):7–16.

[88] Stage TB, Christensen MM, Feddersen S, Beck-Nielsen H, Brosen K. The role of genetic variants in CYP2C8, LPIN1, PPARGC1A and PPARgamma on the trough steady-state plasma concentrations of rosiglitazone and on glycosylated haemoglobin A1c in type 2 diabetes. Pharmacogenet Genomics 2013;23(4):219–27.

[89] Yeo CW, Lee SJ, Lee SS, Bae SK, Kim EY, Shon JH, et al. Discovery of a novel allelic variant of CYP2C8, CYP2C8*11, in Asian populations and its clinical effect on the rosiglitazone disposition *in vivo*. Drug Metab Dispos 2011;39(4):711–6.

[90] Bachmakov I, Glaeser H, Fromm MF, Konig J. Interaction of oral antidiabetic drugs with hepatic uptake transporters: focus on organic anion transporting polypeptides and organic cation transporter 1. Diabetes 2008;57(6):1463–9.

[91] Kalliokoski A, Neuvonen M, Neuvonen PJ, Niemi M. No significant effect of SLCO1B1 polymorphism on the pharmacokinetics of rosiglitazone and pioglitazone. Br J Clin Pharmacol 2008;65(1):78–86.

[92] Stumvoll M, Haring H. The peroxisome proliferator-activated receptor-gamma2 Pro12Ala polymorphism. Diabetes 2002;51(8):2341–7.

[93] Ramirez-Salazar M, Perez-Luque E, Fajardo-Araujo M, Garza SM, Malacara JM. Effect of the Pro12Ala polymorphism of the PPAR gamma 2 gene on response to pioglitazone treatment in menopausal women. Menopause 2008;15(6):1151–6.

[94] Hsieh MC, Lin KD, Tien KJ, Tu ST, Hsiao JY, Chang SJ, et al. Common polymorphisms of the peroxisome proliferator-activated receptor-gamma (Pro12Ala) and peroxisome proliferator-activated receptor-gamma coactivator-1 (Gly482Ser) and the response to pioglitazone in Chinese patients with type 2 diabetes mellitus. Metabolism 2010;59(8):1139–44.

[95] Namvaran F, Azarpira N, Rahimi-Moghaddam P, Dabbaghmanesh MH. Polymorphism of peroxisome proliferator-activated receptor gamma (PPARgamma) Pro12Ala in the Iranian population: relation with insulin resistance and response to treatment with pioglitazone in type 2 diabetes. Eur J Pharmacol 2011;671(1–3):1–6.

[96] Pei Q, Huang Q, Yang GP, Zhao YC, Yin JY, Song M, et al. PPAR-gamma2 and PTPRD gene polymorphisms influence type 2 diabetes patients' response to pioglitazone in China. Acta Pharmacol Sin 2013;34(2):255–61.

[97] Mathur SK, Rathore R, Chandra S, Sharma BB, Sharma L, Mathur P, et al. Primary pioglitazone failure in Asian Indian diabetics is not related to common Pro12Ala polymorph of PPAR-gamma gene. Indian J Physiol Pharmacol 2009;53(2):175–80.

[98] Kang ES, Park SY, Kim HJ, Kim CS, Ahn CW, Cha BS, et al. Effects of Pro12Ala polymorphism of peroxisome proliferator-activated receptor gamma2 gene on rosiglitazone response in type 2 diabetes. Clin Pharmacol Ther 2005;78(2):202–8.

[99] Wolford JK, Yeatts KA, Dhanjal SK, Black MH, Xiang AH, Buchanan TA, et al. Sequence variation in PPARG may underlie differential response to troglitazone. Diabetes 2005;54(11):3319–25.

[100] Florez JC, Jablonski KA, Sun MW, Bayley N, Kahn SE, Shamoon H, et al. Effects of the type 2 diabetes-associated PPARG P12A polymorphism on progression to diabetes and response to troglitazone. J Clin Endocrinol Metab 2007;92(4):1502–9.

This page is intentionally left blank

Chapter 31

Genomics and Pharmacogenomics of Lipid-Lowering Therapies

Kieran F. Docherty and Sandosh Padmanabhan
BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, UK

Chapter Outline

31.1 Overview 715
31.2 Lipid Metabolism 716
  31.2.1 Cholesterol Synthesis 716
  31.2.2 Lipoproteins 717
  31.2.3 Exogenous Pathway 717
  31.2.4 Endogenous Pathway 718
  31.2.5 Reverse Cholesterol Transport 718
31.3 Lipids and CVD Risk, Epidemiology, and Public Health Impact 718
  31.3.1 Total Cholesterol and Low-Density Lipoprotein Cholesterol 719
  31.3.2 HDL-Cholesterol 721
31.4 Genetic Basis of Lipid Disorders and Lipid Levels 722
  31.4.1 Primary Dyslipidemias 722
  31.4.2 Common Genetic Variants and Lipid Levels 727
  31.4.3 Pharmacological Management of Lipids and Pharmacogenomics 728
  31.4.4 HDL-C 738
31.5 Conclusion and Future Directions 741
References 742

31.1 OVERVIEW

Lipids are molecules integral to the body's structure and important substrates for energy metabolism. All the principal blood lipid fractions (total cholesterol, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglycerides (TG)) have both environmental and genetic determinants, with heritabilities of 40–70% [1]. The strong, independent, continuous, and graded relationship between total cholesterol levels, or low-density lipoprotein level, and the risk of coronary artery disease events has been clearly demonstrated worldwide in men and women in all age groups. A causal role for cholesterol was established from observations of individuals with the rare familial hypercholesterolemia (FH), who develop cardiovascular disease (CVD) before the age of 30 in the homozygous state, and from substantive randomized control trials of cholesterol-lowering drug (statins), which showed benefit in terms of reduced CVD and total mortality. The mechanism by which cholesterol increases cardiovascular risk is the process of atherogenesis.

The public health importance of cardiovascular disease is reflected in the Global Burden of Disease survey projections showing that over the next 20 years, the leading cause of death worldwide will be ischemic heart disease and cerebrovascular disease [2]. The mainstay of cardiovascular risk reduction strategies is lipid-lowering therapy, the overall aim of which is to reduce circulating levels of atherogenic lipids in order to slow, reverse, or ideally prevent the development of atherosclerotic plaques. However, like any other drug treatment, the magnitude of plasma lipid responses to drug therapies is subject to considerable interindividual variation. This variation may be due to the multiple genetic factors that affect the pharmacokinetics (absorption, distribution, metabolism, and excretion) or pharmacodynamics (drug targets) of lipid-lowering therapy as well as a multitude of environmental factors. Understanding the basis of

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00031-1
© 2014 Elsevier Inc. All rights reserved.

interindividual variability in drug response will enable the development of biomarkers that will tailor lipid-lowering therapy to maximize efficacy and reduce toxicity.

In this chapter, an overview of lipid metabolism and the genetic basis of dyslipidemia is provided, along with a detailed review of pharmacogenomic advances in lipid-lowering therapy.

31.2 LIPID METABOLISM

Lipids serve as an important energy substrate, building block, and signaling molecule for the human body. They constitute a group of molecules that include sterols, fatty acids, triglycerides, fat-soluble vitamins (A, D, E, and K), and phospholipids among others. Because these molecules are hydrophobic, they must be packaged into particles known as lipoproteins—macromolecular complexes that play a vital role in the absorption and transport of lipids to and from the liver and peripheral tissues.

31.2.1 Cholesterol Synthesis

Cholesterol is an integral component of cell membrane structure and serves as a precursor for bile acids, steroid hormones, and vitamin D. It can be obtained from the diet, but is also endogenously produced. Endogenous production of cholesterol takes place mainly in the liver, with the intestine, adrenal cortex, and gonads contributing to a lesser degree.

The endogenous production of cholesterol, resulting in around 1 gram of cholesterol per day, is a 37-step process with acetyl-CoA as the initial building block. Acetyl-CoA is converted to mevalonate in a three-step process, with the action of HMG-CoA reductase as the rate-limiting step. Three molecules of mevalonate are then decarboxylated to form isopentenyl pyrophosphate, three molecules of which are subsequently condensed by the action of farnesyl synthase to form the intermediary geranyl pyrophosphate and then farnesyl pyrophosphate. The condensation of two molecules of farnesyl pyrophosphate takes place in the endoplasmic reticulum and is catalyzed by squalene synthase to form squalene. The oxidation of squalane by oxidosqualene cyclase, which also occurs in the endoplasmic reticulum, forms lanosterol. Lanosterol undergoes further reactions to produce the end product of a molecule of cholesterol. This process is detailed in Figure 31.1 and was a natural target in the search for drugs to reduce plasma cholesterol concentrations.

Endogenous cholesterol production is under the direct influence of the amount of dietary cholesterol absorbed, with a higher dietary intake resulting in decreased endogenous production. Intracellular cholesterol concentration is

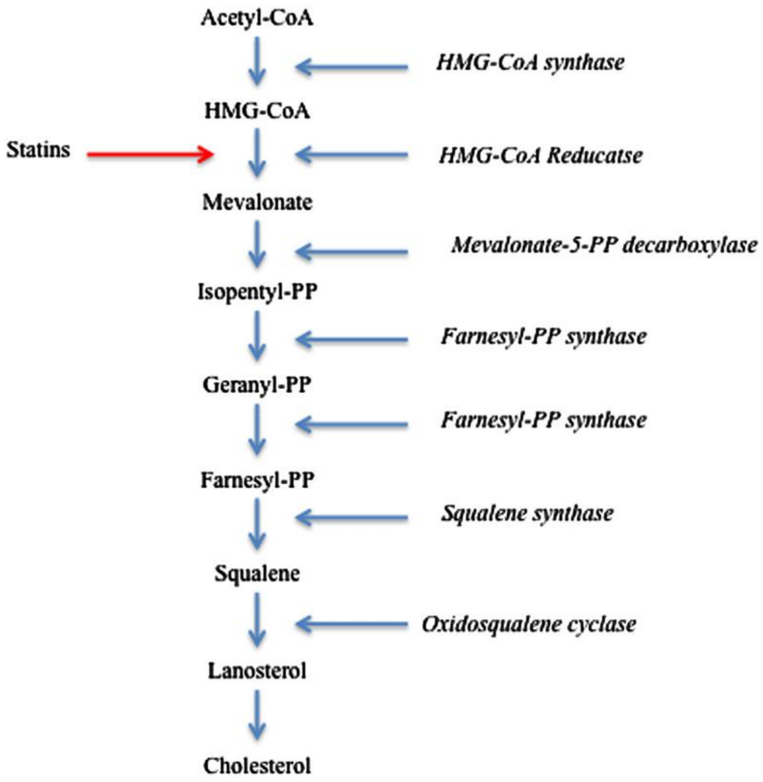


FIGURE 31.1 **Biosynthesis of cholesterol.** Statins act through inhibition of the rate-limiting step, the conversion of HMG-CoA to mevalonate by HMG-CoA reductase.

the main driver for the sterol regulatory element-binding protein (SREBP) pathway. SREBP is a transcription factor bound to the endoplasmic reticulum membrane, which when activated acts on promoter regions of the LDL receptor gene, the HMG-CoA reductase gene, and the HMG-CoA synthase gene. The end result of transcription of these target genes is to upregulate endogenous cholesterol production. Low intracellular cholesterol levels result in activation of the SREBP cleavage-activating protein (SCAP). In times of high intracellular cholesterol concentration, the SREBP-SCAP complex is reversibly bound to the membrane, along with a membrane protein, INSIG1. A decrease in intracellular cholesterol concentration allows dissociation of INSIG1 and the SREBP-SCAP complex. SCAP acts to release SREBP from the endoplasmic reticulum to allow it to travel to the nucleus by the proteolytic cleavage of its two sections by site-1 protease (S1P) and site-2 protease (S2P). When free to travel to the nucleus SREBP acts in the manner described earlier to upregulate endogenous cholesterol production.

31.2.2 Lipoproteins

Plasma lipoproteins consist of a hydrophobic core of cholesterol esters and triglycerides surrounded by an outer hydrophilic layer composed of ampipathic phospholipids, unesterified cholesterol, and apolipoproteins. The lipoproteins are categorized into five classes in a continuum of hydrated density. These are, in ascending order of density: chylomicrons, very-low-density lipoproteins (VLDLs), intermediate-density lipoproteins (IDLs), low-density lipoproteins (LDLs), and high-density lipoproteins (HDLs). Lipoprotein density is dependent on the concentration of triglycerides in the particle. Chylomicrons and VLDL, the largest and least dense lipoprotein particles, are responsible for the majority of the body's triglyceride transport. LDL and HDL are mainly responsible for the transport of the body's cholesterol, with very little of the triglycerides constituent in these particles.

The protein component of lipoprotein particles, apolipoproteins, is integral to the interaction of lipoproteins with cellular receptors, along with the activation and inhibition of enzymes involved in lipoprotein metabolism. Particular apolipoproteins are unique to certain lipoprotein classes. The main apolipoproteins and their characteristics are detailed in Table 31.1.

31.2.3 Exogenous Pathway

Dietary lipids are a vital source of energy to the human body. Triglycerides, which account for more than 95% of dietary lipids, are hydrolyzed by gastric and pancreatic lipases into monoacylglycerol, free fatty acids, and free glycerol. They are then emulsified with bile acids to form micelles, which transport them into the enterocytes of the intestine. The constituents are then repackaged into triglycerides and together with phospholipids, cholesteryl esters, rentinyl esters, cholesterol, and apo B-48 form chylomicrons. This process is driven by the microsomal triglyceride transfer protein (MTP). The newly formed chylomicrons are excreted by the enterocytes into the intestinal lymph system and into the plasma via the thoracic duct.

Chylomicrons travel through the systemic circulation and are significantly metabolized in the peripheral tissues, such as muscle and adipose tissue, before reaching the hepatic circulation and the liver. Circulating chylomicrons are hydrolyzed by lipoprotein lipase (LPL), which

TABLE 31.1 Apolipoproteins and Their Function

Apolipoprotein	Location	Activity
Apo A-I	HDL	Cofactor for LCAT
Apo A-II	HDL	Cofactor for LCAT
Apo A-IV	Chylomicrons	Activates LCAT/CETP
Apo A-V	Chylomicrons, VLDL, HDL	Activates LPL
Apo B-48	Chylomicrons	Ligand for LDLR
Apo B-100	VLDL, IDL, LDL	Ligand for LDLR
Apo C-I	HDL, VLDL	Inhibits CETP
Apo C-II	Chylomicrons, VLDL, HDL	Activates LPL
Apo C-III	Chylomicrons, VLDL, HDL	Inhibits LPL and HL
Apo E	Remnant particles (chylomicrons, VLDL, HDL)	Binds to LDLR

LCAT—lecithin-cholesterol acyltransferase; CETP—cholesterol ester transfer protein; LPL—lipoprotein lipase; HL—hepatic lipase; LDLR—LDL receptor.

is present on the endothelial surfaces of adipose tissue, heart, and skeletal muscle. The hydrolyzation of triglycerides into free fatty acids requires the activation of LPL by apo C-II, which is transferred from HDL to chylomicrons. The free fatty acids are then utilized by the surrounding tissues to generate energy through oxidative reactions, or they are stored as triglycerides.

The resulting chylomicron remnants are rich in cholesterol after the hydrolyzation of the hydrophobic triglyceride components, but also contain apo B-48 and apo E. Apo E is a necessary component in the uptake of chylomicron remnants into the liver by the LDL receptor (also known as the apoB/E receptor) or the LDL receptor-related protein (LRP).

31.2.4 Endogenous Pathway

In the liver, the esterification of fatty acids generates endogenously synthesized triglycerides. These are packaged, along with apo B-100, cholesterol, phospholipids, and vitamin E, to form VLDL particles in a process catalyzed by microsomal triglyceride protein (MTP). VLDL is then secreted into the plasma and undergoes a similar degradation to chylomicrons, driven by the action of LPL in adipose tissue, heart, and skeletal muscle. Prior to this, VLDL particles acquire apo E and apo C-II from HDL particles circulating in the plasma. VLDL particles contain an approximate 1:5 ratio of cholesterol to triglycerides, and subsequent to the action of LPL the resultant IDL particles contain similar amounts of cholesterol and triglycerides.

The liver takes up IDL particles by endocytosis driven by the binding of apo E to the LDL receptor. This accounts for around 40–60% of IDL particles, with the remainder forming LDL by the action of hepatic lipase (HL). The resulting LDL molecules retain apo B-100 and are the most cholesterol-rich of all the lipoproteins.

The majority of circulating LDL is cleared by the binding of apo B and the LDL receptor, resulting in endocytosis into the liver. The remainder is taken up in the peripheral tissues by a similar process. Once in the cells, LDL is degraded by lysosomes and the free cholesterol is esterified by sterol O-acyltransferase (SOAT); this process regulates the body's level of cholesterol through the SREBP pathway described previously.

31.2.5 Reverse Cholesterol Transport

The body cannot metabolize cholesterol, so any excess must be excreted as free sterol molecules via the liver or as bile acids in the gut. The system that drives this process is known as "reverse cholesterol transport" and is mediated by HDL molecules. This process is thought to reduce the risk of atherosclerotic cardiovascular disease by the removal of excess cholesterol in the peripheral tissues to the liver for excretion. (See Figure 31.2.)

HDL particles are synthesized in the intestine and liver as discoidal lipid-poor particles containing apo A-1, also known as pre-β1-HDL. During particle formation, apo A1 interacts with potentially atherogenic cholesterol containing cells expressing the membrane ATP-binding cassette protein A1 (ABCA1), leading to the nascent HDL particles acquiring phospholipids and free cholesterol to form pre-β2-HDL. Once circulating in the plasma, the HDL-associated enzyme lecithin-cholesterol acyltransferase (LCAT) esterifies the free cholesterol; the newly formed hydrophobic cholesterol esters then move to the core of the now mature HDL particle (α-HDL), making it spherical in shape. Further modeling of the HDL particles takes place in the plasma through the action of a variety of cell-surface and plasma enzymes, including hepatic (LIPC) and endothelial lipase (LIPG), the ABCG1 transporter, and the phospholipid transfer protein (PLTP).

Mature HDL particles can be removed from the circulation or transported to steroid-generating organs (adrenals, ovaries, and testes) through a direct or indirect pathway. Direct uptake is mediated by the cell-surface receptor known as scavenger receptor class B1 (SR-B1), which allows cellular uptake of cholesterol from mature HDL particles to form HDL particles that are free to reenter the cycle. The indirect pathway is mediated by cholesterol ester transfer protein (CETP), a plasma protein that facilitates the transfer of cholesterol esters from HDL to apo B-containing lipoproteins (VLDL, LDL) in exchange for a triglyceride particle. This allows cholesterol to be hepatically removed from the circulation through the LDL-R-mediated endocytosis of apo B-containing lipoproteins.

31.3 LIPIDS AND CVD RISK, EPIDEMIOLOGY, AND PUBLIC HEALTH IMPACT

Cardiovascular disease (CVD) is the leading cause of death worldwide and accounts for approximately 17 million deaths yearly—around 30% of total deaths [3]. It is predicted that by 2030 more than 23 million people will die from CVD annually. CVD results from the formation of atherosclerotic plaques in the arteries, leading to morbidity and mortality from coronary artery disease (CAD), peripheral vascular disease (PVD), and ischemic stroke. (See Figure 31.3.)

There are multiple risk factors for the development of CVD; they include modifiable factors—smoking, obesity, lack of exercise, diet, hypertension, diabetes mellitus, and plasma lipid abnormalities—as well as nonmodifiable factors—male gender and age. The relationship between plasma lipid abnormalities and CVD risk has been heavily investigated over the last 50 years, with large advances made in the estimation of cardiovascular risk and the management of dyslipidemias in a primary and secondary prevention setting. In the following section, we explore the individual plasma lipid parameters and their relationship to cardiovascular risk.

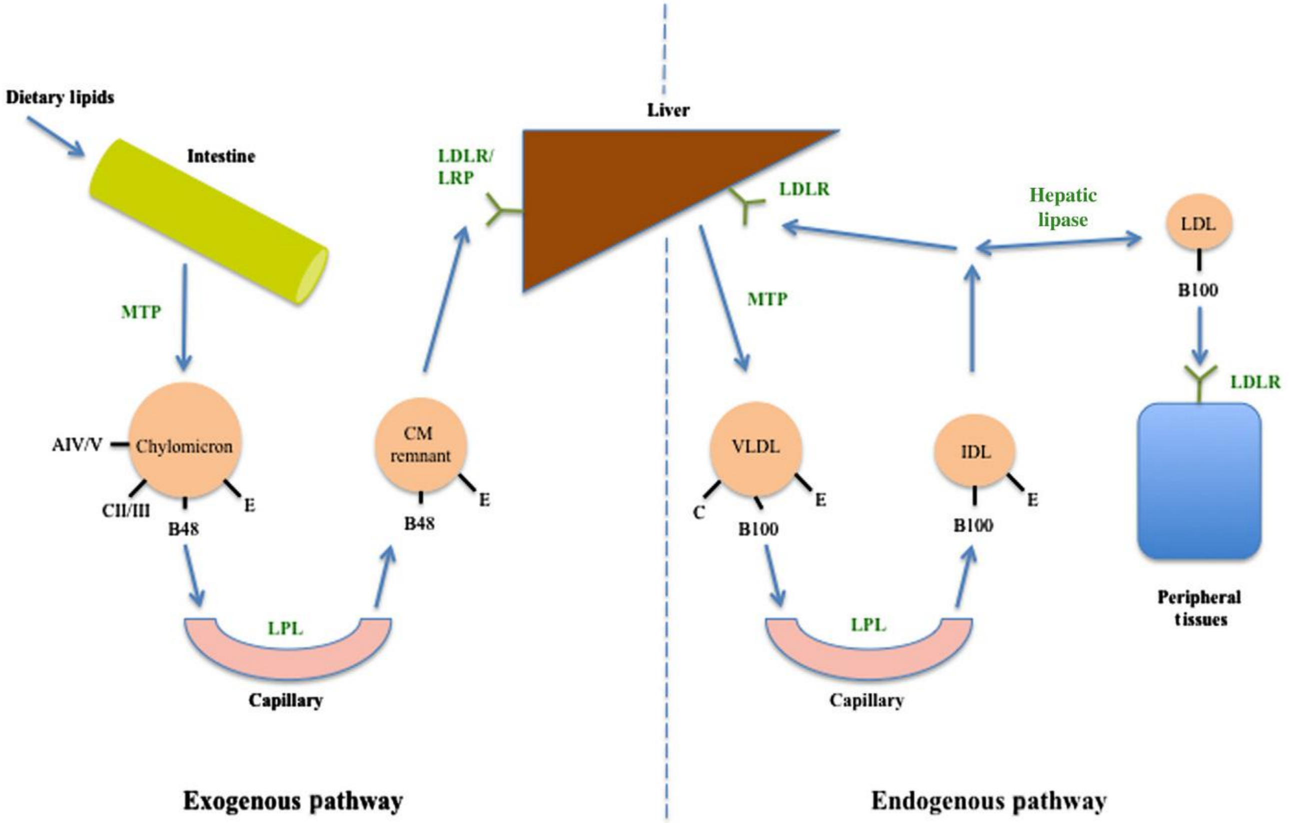


FIGURE 31.2 **Endogenous and exogenous pathways of cholesterol metabolism.** MTP—microsomal triglyceride protein; LPL—lipoprotein lipase; CM—chylomicron; LDLR—LDL receptor; LRP—LDL receptor-related protein; VLDL—very-low-density lipoprotein; IDL—intermediate-density lipoprotein; LDL—low-density lipoprotein.

31.3.1 Total Cholesterol and Low-Density Lipoprotein Cholesterol

The Framingham Heart study (1961) was one of the first observational trials to demonstrate a direct relationship between serum levels of total cholesterol and the development of cardiovascular disease [4,5]. The risk of mortality from coronary heart disease rose with increasing total plasma cholesterol, and these results were echoed in further observational trials, such as the Whitehall study, the Multiple Risk Factor Intervention Trial (MRFIT), and the INTERHEART study.

To investigate the relationship between blood pressure and cholesterol on cardiovascular mortality, a meta-analysis published in 2007 examined 61 prospective observational studies involving 892,000 subjects, an average follow-up of 13 years, and more than 55,000 deaths from a vascular cause [6]. It was demonstrated that a 1 mmol/l reduction in total plasma cholesterol in the age groups of 40–49, 50–69, and 70–89 years was associated with a significant reduction in cardiovascular events; hazard ratios were 0.44, 0.66, and 0.83, respectively.

The individual components of plasma cholesterol, such as LDL-C, HDL-C, and triglycerides, have also been examined in terms of their unique relationship to cardiovascular risk. It was discovered that increasing plasma levels of LDL-C, the major component of the plasma total cholesterol level, increases CVD risk.

As previously noted, elevated levels of total cholesterol and LDL-C are associated with an increased risk of CVD. It therefore follows that reducing levels of these lipid parameters should result in a lowering of this risk. The Scandinavian Simvastatin Survival Study (4S), a landmark trial published in 1994, heralded the role of HMG-CoA reductase inhibitors (statins) in lowering the risk of morbidity and mortality from coronary arterial disease.

A secondary prevention study, involving 4444 patients with established coronary arterial disease and a total cholesterol level of between 5.5 and 8.0 mmol/l, compared the effect of simvastatin versus placebo with a median follow-up of 5.4 years [7]. In the simvastatin arm of the trial, total cholesterol and LDL-C were reduced by 25% and 35%,

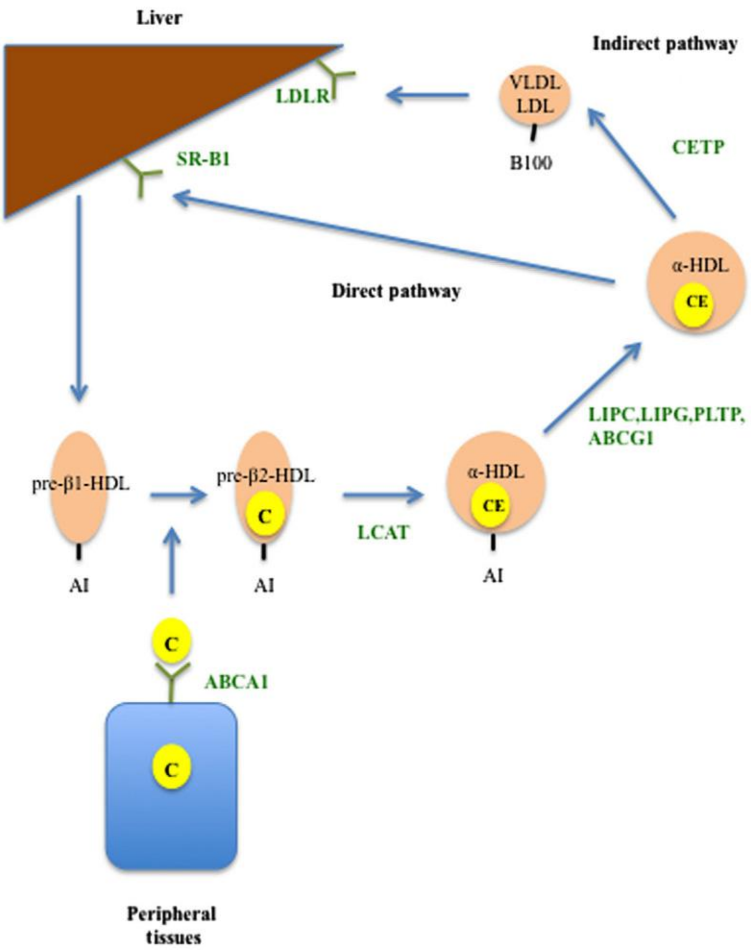


FIGURE 31.3 **Reverse cholesterol transport.** ABCA1—ATP-binding cassette protein A1; LCAT—lecithin-cholesterol acyltransferase; LIPC—hepatic lipase; LIPG—endothelial lipase; ABCG1—ATP-binding cassette protein G1; PLTP—phospholipid transfer protein; CETP—cholesterol ester transfer protein; LDLR—LDL receptor; SR-B1—scavenger receptor class BI.

respectively. This resulted in a statistically significant relative risk reduction of 0.66 for major coronary events (coronary death, myocardial infarction, or resuscitated cardiac arrest), and a reduced total mortality with a relative risk reduction of 0.70. These results indicated no difference in noncardiovascular death in the treatment and placebo groups.

The role of statins in secondary prevention was further confirmed by the results of the Cholesterol and Recurrent Events (CARE) trial, in which patients with a history of myocardial infarction and lower cholesterol levels than those in 4S (<6.2 mmol/l) showed a significant reduction in the incidence of coronary events compared to placebo [8].

The West of Scotland Coronary Prevention Study (WOSCOPS) trial was the first primary prevention trial to examine the role of plasma cholesterol reduction in hypercholesterolemic men with no history of coronary heart disease [9]. This was a high-risk population with average total plasma cholesterol levels of 7.03 mmol/l, one-third of the 6595 study group being smokers with an average body mass index (BMI) of 26. Participants were randomized to either pravastatin 40 mg once daily or placebo. In the treatment group, total cholesterol and LDL-C were reduced by 20% and 26%, respectively. The lowering of plasma cholesterol levels resulted in a reduced rate of coronary events by 31% and cardiovascular mortality by 32%. This study illustrated that statin-induced hypercholesterolaemia lowering, even in the absence of known coronary heart disease, had a positive effect on the rate of cardiovascular events and mortality.

The individual trials just discussed demonstrate a clinical benefit to the lowering of total plasma cholesterol and LDL-C as it relates to cardiovascular risk. In addition, there appears to be no threshold under which any further decrease in LDL-C does not result in lowered risk. These findings have been confirmed by a large meta-analysis by the Cholesterol Treatment Trialists' (CTT) Collaboration, where data from more than 90,000 patients with more than 14,000 vascular events were analyzed [10]. It was demonstrated that statin therapy, resulting in a lowering of LDL-C, reduces the risk of vascular events and death [8]. Each reduction in LDL-C by 1 mmol/L resulted in a reduction in mortality from CVD by around one-fifth. Further lowering also equated to further risk reduction. The results also suggest that the benefit incurred from reduction of LDL-C is the same irrespective of pretreatment LDL-C level. That is to say, benefit is accrued from the absolute reduction in LDL-C level, not from the particular level reached. The CTT meta-analysis included patients who had a history of CVD and those who did not.

Another meta-analysis, published in 2013, has shown benefit in LDL-C reduction with no excess of adverse effects in a selected patient subgroup with no history of CVD—that is, a group in which the risk of CVD is lower [11]. This finding echoed those of the CTT collaboration—that is, similar reduction of absolute risk by lowering LDL-C in patients at low risk of CVD and those at higher risk.

In conclusion, elevated plasma levels of total cholesterol and LDL-C are known to increase CVD risk; reducing those levels reduces that risk in both primary and secondary prevention settings.

31.3.2 HDL-Cholesterol

Lower LDL-C, as noted earlier, results in a lower cardiovascular risk. However, despite aggressive lowering, there still exists residual CVD risk. Data from the Framingham Heart Study and subsequent studies have shown low plasma levels of HDL-C to be an independent risk factor for CVD [12,13]. In the Framingham study, HDL-C levels were a more potent risk factor for CVD than plasma levels of LDL-C. In the Treating to New Targets (TNT) trial, where patients with known CAD were treated with statins to achieve an LDL-C level of <1.8 mmol/L, HDL-C level remained an independent risk factor when adjusted for other baseline risk factors [14]. A meta-analysis of four trials, adjusting for age, blood pressure, smoking, BMI, and LDL-C, found that an increase in HDL-C of 1 mg/dL (0.03 mmol/L) resulted in a decrease of 1.9–2.9% in risk [15]. These results suggest that increasing low levels of HDL-C should result in decreased CVD risk. However, there is limited data to support this hypothesis in trials using pharmacological methods to raise HDL-C using drugs such as nicotinic acid and CETP inhibitors. Given these disparate findings, the exact role of HDL-C as an independent factor in determining cardiovascular risk is still under debate.

31.3.2.1 Triglycerides

The potential association between elevated triglyceride levels and cardiovascular disease has been investigated for more than fifty years, with no conclusive evidence of an independent relationship as a risk factor. Several prospective studies have shown elevated plasma triglyceride levels to be a CVD risk factor. An analysis carried out on two prospective cohort studies, the Reykjavik study and the European Prospective Investigation of Cancer (EPIC) Norfolk study, along with a meta-analysis of 29 studies involving more than 10,000 CVD events in 262,525 individuals, calculated a combined odds ratio after adjustment for established CVD risk factors of around 1.7 in subjects with a triglyceride level in the highest tertile [16].

This finding, in a Western population, was very similar to that found in a previous analysis of an Asian-Pacific population [17]. A further meta-analysis in 2009 of more than 300,000 people, with 2.79 million person-years of follow-up, showed that the hazard ratio for CVD per 1 standard deviation higher triglyceride plasma concentration was 1.37 (95% CI 1.31–1.42) [18]. However, after adjustment for values of well-established risk factors, including non-HDL and HDL cholesterol, the hazard ratio was reduced to 0.99 (95% CI 0.94–1.05).

The results just discussed suggest a possible association between elevated triglyceride levels and CVD risk, which lessens after adjustment for well-established CVD risk factors, in particular plasma levels of HDL-C and LDL-C. A study published in 2010 aimed to clarify the role of triglyceride-mediated pathways in promoting CVD risk [19]. One of the most common genetic variants associated with triglyceride levels, rs662799 (−1131T>C) in the apolipoprotein A-5 (APOA5) gene, codes for the apolipoprotein associated with VLDL, a molecule rich in triglycerides. The authors examined the association between this genetic variant, elevated triglyceride levels, and the risk of coronary heart disease. They found that, for each minor C allele inherited, the mean triglyceride level was higher by 16% (0.25 mmol/L). There was no significant association between rs662799 and LDL-C concentration, and only a modest association with lower levels of HDL-C (mean difference of 3.5% per C allele inherited), lower apolipoprotein A1 (1.3%), and higher apolipoprotein B (3.2%).

Analysis findings equated to an odds ratio of coronary heart disease of 1.18 (95% CI 1.11–1.26) per inherited C allele. An additional finding was that the rs662799 polymorphism, and thus hypertriglyceridemia, was associated with higher VLDL plasma concentration and smaller HDL-C particles, thought to be potential pro-atherogenic precipitants of increased CVD risk. The authors concluded that there was a causal association between triglyceride-mediated pathways and coronary heart disease.

31.3.2.2 Apolipoproteins

Outside of the now standard measurements of LDL-C, HDL-C, and triglycerides in the estimation of cardiovascular disease risk, recent research has pointed toward the measurement of plasma apolipoprotein concentrations as potentially more accurate markers of CVD risk.

Apo B/Apo A-1

Apolipoprotein B is the major constituent of the atherogenic particles VLDL, IDL, and LDL. The plasma concentration of apolipoprotein B can be measured and has been examined for its function as a surrogate marker for cardiovascular risk. In addition to apo B, non-high-density lipoprotein cholesterol, the sum of VLDL, IDL, and LDL particles in the plasma (calculated as total cholesterol minus HDL-C) has been used in a similar manner [20].

Apolipoprotein A-1 is the major constituent of HDL-C and acts as a surrogate marker for HDL-C plasma concentration. Several prospective studies have shown measurement of apo-B or non-HDL-C to be at least as good as measurement of LDL-C in CVD risk reduction. The Emerging Risk Factors Collaboration (ERFC) found that CVD hazard ratios for non-HDL-C and HDL-C were almost identical to those found for apo B and apo A-1 in a meta-analysis of more than 300,000 subjects [18]. A later analysis, including the ERFC study set and totaling 233,455 subjects and 22,950 cardiovascular events, calculated apo B concentration to be a more powerful predictor of risk than non-HDL-C and LDL-C, with relative risk ratios of 1.43 (95% CI 1.35–1.51), 1.34 (95% CI 1.24–1.44), and 1.25 (95% CI 1.18–1.33), respectively [21].

With these results in mind, the authors calculated that measurement of apo B to predict risk and guide management could potentially, over a 10-year period, prevent 500,000 more cardiovascular events than a strategy based on non-HDL-C measurement. In comparison, the measurement of non-HDL-C was calculated to prevent 300,000 more events than a LDL-C measurement strategy.

Lipoprotein(a)

Lipoprotein(a) (Lp(a)) consists of an apolipoprotein B-100 particle attached to apolipoprotein A by a disulphide bridge. Recent research has shown elevated plasma concentrations of lipoprotein(a), in particular the apo A component, to be atherogenic, leading to increased levels of coronary artery disease. In an analysis of 36 prospective studies with more than 120,000 participants and 22,000 vascular events, there was a risk ratio of 1.13 (95% CI 1.09–1.18) per 1 standard deviation increase in Lp(a) for CVD, after adjustment for well-established CVD risk factors [22]. Genetic variants in Lp(a) have been identified, placing subjects with the genotypes at a higher risk; rs3798220 and rs10455872, which are associated with elevated plasma levels of Lp(a) and increased risk of CVD, are two SNPs that have been identified [23]. Current guidance does not recommend routine measurement of Lp(a) as part of risk estimation, but it may be considered in people with a high risk or strong family history of CVD [24].

31.4 GENETIC BASIS OF LIPID DISORDERS AND LIPID LEVELS

As detailed previously, the associations between plasma lipid abnormalities and increased risk of cardiovascular disease are well established. Recent advances in determining the genetic basis for plasma lipid levels have led to new insight into the pathogenetic basis of increased cardiovascular risk in certain individuals. Cardiovascular risk is a complex heterogeneous issue with no single etiological factor. It is well established that cardiovascular disease has a strong element of heritability, with a positive history in previous generations being a strong independent risk factor. However, the precise mode of inheritance is still unclear. In their individual components, lipid levels have been shown to have an element of heritability, with levels of correlation seen between family members against nonrelated individuals.

The association between lipid abnormalities and increased cardiovascular risk has long been established in Mendelian disorders, such as the primary dyslipidemias described next. The mapping of the human genome has led to the advent of genome-wide association studies (GWASs), with the aim of identifying common genetic variants that associate with lipid levels and consequently cardiovascular disease. The identification of functional genomic variants, which predispose to increased risk, has the potential to allow presymptomatic identification of those at high risk earlier than current practice, and allows early-targeted intervention to reduce risk. This is discussed later in the chapter.

31.4.1 Primary Dyslipidemias

Several monogenic disorders result in abnormalities of an individual's lipid profile, with varying corresponding levels of CVD risk. Table 31.2 outlines the most common of these disorders, their genetic basis, and their biochemical effect, along with their prevalence and distinguishing clinical features. We examine disorders affecting each aspect of the lipid profile in turn.

31.4.1.1 Primary LDL Abnormalities

Familial Hypercholesterolemia

The molecular defect causing familial hypercholesterolemia (FH) was first identified by Brown and Goldstein in their Nobel prize-winning work published in 1986 [25]. This discovery elucidated the role of LDL receptors in the uptake of cholesterol and the loss-of-function mutation in the LDLR gene located on chromosome 19p13.1–13.3, leading to the elevated LDL-C levels seen in FH. Two genotypic variants with an autosomal dominant pattern of inheritance exist, and these have differing clinical and biochemical presentations [26].

TABLE 31.2 Primary Monogenic Dyslipidemias

Genetic Basis Mode of Inheritance, Prevalence, and Distinguishing Clinical Features

Disorder	Genetic Defect	Lipid Defect	Inheritance	Prevalence	Clinical Findings
LDL-C					
Familial hypercholesterolemia	LDLR	↑ LDL	ACD	Heterozygous: 1/500	Heterozygous (TC: 7–13 mmol/L): tendon xanthomata, arcus senilis, premature CAD (age 30–50)
				Homozygous: 1/1,000,000	Homozygous (TC >13 mmol/L): plantar/tendon xanthomata, premature CAD (age <18)
Familial defective apoB-100	APOB (LDL receptor binding-region defect)	↑ LDL	AD	1/700	TC 7–13 mmol/L: tendon xanthomas, arcus senilis, premature CAD
Autosomal recessive hypercholesterolemia	LDLRAP1 (loss-of-function)	↑ LDL	AR	Very rare	Tendon xanthomas, arcus senilis, premature CAD
Autosomal dominant hypercholesterolemia 3	PCSK9 (gain-of-function)	↑ LDL	AD	Rare	Tendon xanthomas, arcus senilis, premature CAD
Sitosterolemia	ABCG5/8	↑ LDL	AR	Rare	Tendon xanthomas, premature CAD
Abetalipoproteinemia (Bassen-Kornzweig syndrome)	MTP (microsomal triglyceride transfer protein), Apo B-48/B-100	↓ LDL	AR	Rare	Fat malabsorption, spinocerebellar degeneration, retinopathy, acanthocytosis
Familial hypobetalipoproteinemia	APOB (↓ Apo B100)/PCSK9	↓ LDL	ACD	Very rare	Fat malabsorption, fatty liver, spinocerebellar degeneration, retinopathy
PCSK9 deficiency	PCSK9 (loss-of-function)	↓ LDL	ACD	Rare	Reduced CAD risk
HDL-C					
APOA-1 mutations	APOA1 (deletion/mutations)	↓ HDL	AD	Rare	HDL 15–30 mg/dL: corneal opacities, xanthomas, premature CAD
Familial LCAT deficiency	LCAT	↓ HDL	AR	Very rare	HDL-C <10 mg/dL: corneal opacities, hemolytic anemia, renal failure, hepatosplenomegaly

(*Continued*)

TABLE 31.2 Continued

Disorder	Genetic Defect	Lipid Defect	Inheritance	Prevalence	Clinical Findings
Fisheye disease (partial LCAT deficiency)	LCAT	↓ HDL	AR	Very rare	HDL-C: <10 mg/dL corneal opacities, hepatosplenomegaly
Tangier disease (analphalipoproteinemia)	ABCA1	↓ HDL	AR	Rare	HDL-C: <5 mg/dL premature CAD, peripheral neuropathy, hemolytic anemia, corneal opacities, hepatosplenomegaly, orange tonsils
CETP deficiency	CETP	↑ HDL	AR	Rare	Elevated HDL
Triglycerides					
LPL deficiency	LPL (endothelial LPL defect)	↑ Chylomicron	AR	Rare	TG: >8.5 mmol/L: infantile failure to thrive, eruptive xanthomas, hepatosplenomegaly, pancreatitis
Familial apolipoprotein C-II deficiency	APOCII (functional LPL deficiency)	↑ Chylomicron	AR	1/1,000,000	TG >8.5 mmol/L: pancreatitis, metabolic syndrome, eruptive xanthomas, hepatosplenomegaly
Familial hepatic lipase deficiency	LIPC	↑ VLDL remnants	AR	Very rare	Premature CAD
Familial dysbetalipoproteinemia	APOE	↑ Chylomicron and VLDL remnants	AD	1/10,000	TC 6.5–13.0 mmol/L; TG: 2.8–5.6 mmol/L: palmar/tuberoeruptive xanthomas, premature CAD

Heterozygous FH, where there is one mutated allele, is one of the most common genetic lipid disorders, with a prevalence of around 1 in 500. Affected individuals have elevated total cholesterol levels, typically between 7 and 13 mmol/L, resulting in an increased risk of atherosclerotic disease, which typically affects them starting at age 30.

Homozygous FH involves a mutation in both alleles coding for LDLR, resulting in markedly elevated total cholesterol levels, often more than 13 mmol/L. It is less common than heterozygous FH, with a prevalence of around 1 in a 1,000,000, and individuals are affected by an extremely early onset of atherosclerotic disease, often in adolescence. Given the extremely high LDL-C levels these individuals exhibit, treatment involves the use of high-dose statins and adjuncts such as cholesterol absorption inhibitors, bile acid sequestrants, nicotinic acid, or LDL apheresis in extreme cases.

During the early development of statins, Merck suspended a clinical trial of lovastatin because of concerns regarding potential toxicity; consequently, the future of the drug seemed extremely doubtful. This changed when small-scale studies in selected small groups of patients with severe heterozygous FH refractory to existing therapy showed dramatic reductions in LDL cholesterol with few adverse effects. Merck reinitiated its clinical development program in 1983 [27].

Familial Defective apo B-100

Familial defective apo B-100 is clinically indistinguishable from FH. However, it involves a different genetic basis with a mutation in the receptor-binding region of the apo B-100 gene, located on chromosome 2p24-p23 [28]. Several specific mutations have been identified, with the most common involving the substitution of glutamine for arginine at position 3500 in apo B-100. Elevated LDL-C levels result from the failure of the binding between circulating LDL-C in plasma and LDLR. This genetic defect is not as common as that of the LDLR, with a prevalence of around 1 in 700; it is treated in a similar manner to heterozygous FH.

Autosomal Recessive Hypercholesterolemia

Autosomal recessive hypercholesterolemia (ARH) is a rare cause of elevated LDL-C caused by a loss-of-function mutation in the LDLRAP1 gene, located on chromosome 1p36-35; it codes for the adaptor protein ARH [29,30]. This LDLRAP1 mutation results in the inability of hepatocytes to internalize the LDL-C ligand-receptor complex, resulting in elevated LDL-C plasma levels. Individuals with a heterozygous genotype often have normal cholesterol levels, unlike a homozygous genotype, where the clinical presentation is similar to that of homozygous FH.

Autosomal Dominant Hypercholesterolemia

Autosomal dominant hypercholesterolemia (ADH) can be caused by a defect in one of three genes, two of which were described in detail earlier: LDLR and APOB100. In 2003, a third causative mutation was identified in the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene, located on chromosome 1p34.1-p32 [31]. PCSK9 is a protease enzyme that plays a vital role in cholesterol homeostasis by binding to the LDLR and enhancing its endosomal degradation, resulting in increased LDL catabolism. In the subtype of ADH caused by two possible gain-of-function mutations in the PCSK9 gene (HCHOLA3), there is increased PCSK9 protease-driven degradation of LDLR, resulting in decreased uptake of cholesterol into cells and subsequently elevated plasma levels. The impact of these mutations on cardiovascular disease are discussed later in the chapter, along with the ongoing efforts to establish the potential of PCSK9 as a pharmacological target with the aim of reducing levels of LDL-C and thus CVD risk.

Sitosterolemia

Sitosterolemia is an extremely rare autosomal recessively inherited condition caused by mutations to one of two genes, ABCG5 and ABCG8, located on chromosome 2p21, which encode for two adenosine triphosphate-binding cassette (ABC) transporter proteins, sterolin-1 and sterolin-2 [32]. The mutation leads to increased intestinal absorption and reduced biliary excretion of plant sterols with subsequent elevated plasma levels of sitosterol, which result in reduced hepatic LDLR expression, elevated LDL-C, and an associated increased risk of premature CVD.

Abetalipoproteinemia (Bassen-Kornzweig syndrome)

Abetalipoproteinemia (ABL) is a rare autosomal recessively inherited condition that results from a mutation in the gene located on chromosome 4q24 encoding for the microsomal transfer protein (MTP) [33]. This protein plays an important role in the assembly of lipoproteins in the endoplasmic reticulum. The failure of the transfer of lipids to lipoproteins containing apo B-48 and B-100 results in markedly reduced plasma levels of apo B-containing lipoproteins such as V-LDL and chylomicrons. This gives rise to very low plasma cholesterol levels and presents clinically with fat malabsorption, spinocerebellar degeneration, retinopathy, and acanthocytosis due to poor intestinal absorption of the fat-soluble vitamins A, D, K, and, in particular, E.

Familial Hypobetalipoproteinemia

Individuals with familial hypobetalipoproteinemia (FHBL) have LDL-C levels in the <5th percentile and an associated low risk of CVD. The genetic cause for this condition in the majority of cases is mutations in the APOB gene [34]. Multiple mutations have been identified, the majority of which result in truncated forms of Apo B with reduced plasma levels of apo B-100; these are thought to be inherited in an autosomal codominant manner. Heterozygotes can be asymptomatic or demonstrate fatty liver, intestinal fat malabsorption. Homozygotes present in a manner similar to that in patients with ABL. Some patients with the FHBL phenotype have been found to have no mutation in the *APOB* gene, and patients with the phenotype have been found to harbor mutations in *PCSK9*.

PCSK9 Deficiency

Similar to the gain-of-function mutations described, several loss-of-function mutations in PCSK9 have been discovered in individuals with low LDL-C plasma levels and low CVD risk [35].

31.4.1.2 Primary HDL Abnormalities

APOA-1 Mutations

A variety of genetic mutations to the *APO A-1* gene have been identified that lead to low HDL-C levels and to the development of atherosclerotic disease and premature coronary heart disease [36]. Individuals with a complete loss of apo A-1 have very low levels of HDL-C (<5 mg/dL) and correspondingly high-risk levels of premature CAD. Heterozygotes demonstrate no specific clinical symptoms; however, they do have HDL-C and apo A-1 levels around 50% of normal. Several missense or nonsense mutations of apo A-1 have been identified as causing low HDL-C levels and premature CAD. Thus, apo A-1, thought to be atheroprotective, has become a potential pharmacological target. An exception to these findings is the apo A-1_{Milano}

mutation, which, although causing reduced HDL-C levels, has actually been shown to result in reduced risk of cardiovascular events [37].

Analphalipoproteinemia (Tangier Disease)

Analphalipoproteinemia, or Tangier disease, is an extremely rare condition, resulting in decreased plasma HDL-C levels, that is inherited in an autosomal recessive manner [38,39]. It is caused by a loss-of-function mutation to the ABCA1 gene on chromosome 9q31, which encodes for the ATP-binding cassette transporter (ABCA1). ABCA1 is responsible for regulating the efflux of cholesterol and phospholipids to lipid-free apo A-1, forming nascent HDL-C. The mutation to the ABCA1 gene results in a lack of HDL-C in plasma and a toxic buildup of cholesterol in the cells, leading to impaired cellular function and cell death.

Tangier disease presents clinically with peripheral neuropathy, hemolytic anemia, corneal opacities, hepatosplenomegaly, and orange tonsils due to the accumulation of cholesterol in cells, in particular the reticuloendothelial system. The relationship between Tangier disease and CVD risk has not been fully elucidated, partly because of the rarity of the condition. However, it does not appear to cause as early an onset of CVD as that seen in diseases such as homozygous FH. This may be due to the low plasma levels of LDL-C seen in individuals with Tangier disease.

Familial LCAT Deficiency/Fisheye Disease (Partial LCAT Deficiency)

The lecithin-cholesterol acyltransferase (*LCAT*) gene, located on chromosome 16q22, is responsible for encoding the LCAT enzyme, which catalyzes the formation of cholesterol esters from free cholesterol in nascent HDL. This allows the HDL particles to mature. Homozygosity for loss-of-function mutations in the *LCAT* gene can present in two ways, with heterozygotes having HDL-C levels within the normal range [40]. Individuals with familial LCAT deficiency have a complete deficiency of LCAT and are unable to form mature HDL-C and esterify free plasma cholesterol. This presents clinically as HDL-C levels lower than 10 mg/dL with corneal opacification, hemolytic anemia, and progressive renal impairment, which in some cases results in end-stage renal failure. Fisheye disease is a variant with partial LCAT deficiency, where LCAT is active on LDL but not HDL, resulting in the same clinical picture apart from the fact that there is no effect on renal function. As with Tangier disease, no clear association between LCAT deficiency and CVD has been established.

CETP Deficiency

CETP deficiency is inherited in an autosomal recessive manner, in which affected individuals have loss-of-function mutations in both copies of the *CETP* gene. This results in elevated HDL-C levels and low LDL-C levels due to the failure of CETP-driven transfer of cholesterol esters from HDL to apo B-containing lipoproteins such as VLDL, IDL, and LDL. Several mutations have been identified, and data suggest that this genetic disorder is most prevalent in the Japanese population [41]. The two most common mutations are a G>A at +1 position on intron 14 and a D>G substitution at the 442 position on exon 15 (D442G), which are respectively present in approximately 2% and 7% of the Japanese population. Given that inhibition of CETP activity results in lower production of atherogenic lipoproteins, it has been suggested that CETP deficiency is associated with lower CVD risk [42]. To date, the evidence is conflicting, with some studies suggesting a low risk of CVD associated with CETP deficiency, and some showing equivocal results. This issue is further complicated by the current failure of pharmacological CETP inhibition to have any benefit on cardiovascular outcomes.

31.4.1.3 Primary Triglyceride Abnormalities

Familial Chylomicronemia Syndrome

Familial chylomicronemia syndrome (FCS)—that is, LPL deficiency and apo C-II deficiency—is a clinical syndrome of extreme hypertriglyceridemia presenting with eruptive xanthomas, lipemia retinalis, hepatosplenomegaly, and acute pancreatitis [43]. It is caused by homozygosity for one of two loss-of-function mutations in either the LPL or apo C-II genes. These mutations result in a failure of hydrolyzation of triglycerides by LPL and its cofactor, apo C-II, causing elevated plasma triglyceride chylomicron levels. There is no association between FCS and increased risk of atherosclerotic cardiovascular disease. Management involves reduction of dietary fat intake and avoidance of alcohol.

Familial Dysbetalipoproteinemia

Familial dysbetalipoproteinemia (FD), also known as type III hyperlipoproteinemia, arises from mutations in the gene coding for apo E [44]. APOE, normally expressed as the wild-type APOE3, acts as a ligand for the hepatic uptake of chylomicrons, IDL, and VLDL remnant particles. FD is seen in individuals homozygous for a common variant (rs7412) of APOE, known as APOE2, which results from a substitution of a cysteine for an arginine at position 158 on chromosome 19. The E2/E2 genotype has a prevalence of around 1 in 100. However, FD is only seen in around 1 in 10,000 individuals, suggesting that additional environmental, hormonal (estrogen deficiency), and genetic factors, in addition to the genotype, are required for disease development [45]. Individuals with FD have elevated levels of triglycerides and total cholesterol, and present clinically with palmar and tuberoeruptive xanthomata. An increased prevalence of early-onset CAD is also seen.

Hepatic Lipase Deficiency

Hepatic lipase deficiency is a very rare condition where deficiency in hepatic lipase results in elevated levels of total cholesterol, in particular HDL-C, and triglycerides. Various loss-of-function mutations have been identified in individual groups in the *LIPC* gene, which codes for hepatic lipase [46,47]. Although the rarity of this disease precludes definitive association, it is thought that the lipid profile of affected individuals results in increased CVD risk.

31.4.2 Common Genetic Variants and Lipid Levels

Research on the role of gene variants in determining plasma lipid levels has identified common genetic variants in more than 100 gene loci. The largest meta-analysis of GWASs to date examined 100,184 individuals of European ancestry and identified 95 loci demonstrating genome-wide significance ($p<5E\text{-}8$) with one or more of the traits of total cholesterol (TC), LDL-C, HDL-C, and TG [48]. The authors calculated that the loci identified accounted for 10–12% of total variance and 25–30% of genetic variance. This level of variance points toward the existence of other loci and unknown independent signals at already identified loci.

In analysis of the loci identified, 29 of 95 were associated with coronary artery disease, the majority being associated with elevated LDL-C, which supports the well-established causal relationship between elevated LDL-C and CAD risk. A further study performed in 2012, using a dense gene-centric approach in a meta-analysis of 32 studies (66,240 individuals of European ancestry), discovered additional loci associated with lipid levels [49]. The researchers discovered a further 21 novel loci and similar levels of variance as seen in the Tesslovich et al. study. These loci, taken on an individual basis, have relatively small “effect sizes,” and it is thought likely that an individual’s lipid profile is determined by a combination of environmental factors along with variants in multiple genes.

The question of genetic lipid level variance between different ethnic groups has arisen because of the known ethnic variation in lipid levels generally. This potential genetic basis for variance has been studied with known loci and in the search for novel loci as well. Tesslovich et al. confirmed that the majority of the 95 loci identified in their meta-analysis replicated their effect on lipid traits in South Asian, East Asian, and African-American groups.

The role that these common genetic variants potentially play in elevating CVD risk has also been examined. In an analysis of two prospective cohorts, the Whitehall II and British Women’s Heart and Health Study, investigators calculated lipid genetic scores based on subjects’ expression of SNPs known to be causal variants [50]. They found that individuals with high scores had higher levels of LDL-C, a higher chance to be deemed “high risk” by the Framingham risk calculator, and, most important, a higher odds ratio for cardiovascular events. Neither HDL-C nor triglyceride scores demonstrated a similar finding.

31.4.2.1 LDL-C

Variance in LDL-C levels is known to be due to genetic variants coding for Mendelian disorders, including LDLR, PCSK9, APOB, LDLRAP, and ABCG5/8, as described before. In addition, loci for HMGCR and NPC1L1 are known LDL-C lowering pharmacological targets for statins and ezetimibe, respectively. GWASs, such as that described earlier, have led to multiple novel loci being identified in association with variance in LDL-C levels. These include common variants at APOB, APOE, HMGCR, LDLR, and PCSK9 [48]. The 1p13 locus, coding for SORT1, has been identified as the strongest statistical association with LDL-C and CAD in the entire human genome, and homozygosity for the minor allele of 1p13 has been associated with a 40% reduction in risk of myocardial infarction [51–53].

In comparison, homozygotes of the minor allele have, on average, an LDL-C level that is 16 mg/dL lower than that in homozygotes for the major allele. This is a considerable effect size compared to other common variants. The exact role of SORT1 in LDL metabolism, and subsequent CVD risk, has not been clarified; several studies report disparate results in terms of the direction of effect of SORT1 expression [51,52,54]. Given the obvious pivotal role of SORT1 variance in LDL-C levels, it is hoped that the exact mechanism can be elucidated, and its possible role as a therapeutic target explored in the near future.

31.4.2.2 HDL-C

The heritability for HDL-C has been shown in twin studies to range from 40–60%; however, examination of common variants has shown a heritability of only 10–12% from these variants [55]. This suggests that our knowledge of the genetic basis of HDL-C levels is far from complete; this is in addition to the uncertainty surrounding the role of HDL-C in determining CVD risk. As with LDL-C, several monogenic disorders alter lipid levels. These include APOA1, ABCA1, LCAT, CETP, and LIPG. It should be noted, however, that Mendelian randomization studies of subjects with these genetic disorders have failed to show any association of genetically induced low HDL-C with CVD risk [56]. In addition, GWASs have uncovered several SNPs with an association with HDL-C level, including IRS1, ANKS1, KLF14, LPL, TRIB1, the ZNF259–APOA5–APOA4-APOC3–APOA1 gene cluster, and SH2B3 [57].

31.4.2.3 Triglycerides

Several monogenic determinants of hypertriglyceridemia have been identified, and the more common were explored earlier in the chapter. They include mutations at LPL, APOCII, and APOA5, along with rarer mutations at LMF1 (lipase maturation factor 1) and GPIHBP1. LMF1 is involved in the maturation and transport of lipoprotein lipase. Individuals with hypertriglyceridemia who lack mutations in the more common genes associated with elevated triglyceride levels have later been found to have loss-of-function mutations in LMF1 [58]. GPIHBP1 encodes for glycosylphosphatidylinositol-anchored HDL binding protein, which plays a key role in the processing of chylomicrons mediated by LPL. Homozygous mutations in GPIHBP1 have been identified in subjects with hypertriglyceridemia who lack more common mutations at other associated loci [59].

As with LDL-C and HDL-C, current studies have identified common genetic variants that account for around 10% of total population triglyceride variance, suggesting that our knowledge of the genetic determination of triglyceride levels is far from complete. Beyond monogenic disorders, it is thought that a combination of genetic variants is responsible for the variance seen. Loci most strongly associated with triglyceride levels in GWASs include APOA5, GCKR, LPL, and APOB, alongside other loci identified at TRIB1, ANGPLT3, MLXIPL, and KLHL8 [48,60].

31.4.3 Pharmacological Management of Lipids and Pharmacogenomics

We discussed earlier the role of the dyslipidemias in elevating CVD risk. Various pharmacological methods of managing dyslipidemia have been identified and investigated, with varying results in terms of their ability to reduce this risk. We examine each fraction of the lipid profile in turn and discuss the established and novel pharmaceutical methods available. The principle pharmaceutical agents currently available are detailed in Table 31.3 along with their individual effect on the plasma lipid profile.

31.4.3.1 LDL-C

Statins

As detailed earlier, statin therapy is currently universally accepted as the most effective pharmacological therapy for lowering total plasma cholesterol and LDL-C. A reduction in plasma levels leads to decreased CVD-associated morbidity and mortality. In addition to their effect on total cholesterol and LDL-C, statins, to a lesser extent, reduce triglyceride levels and effect an increase in HDL-C. This section of the chapter examines the pharmacology of statins, common side effects, and how common genetic variants can potentially alter individual patient response to therapy and increase the risk of certain side effects.

Statins are competitive inhibitors of hydroxymethylglutaryl-coenzyme A (HMG-CoA) reductase, the enzyme controlling the rate-limiting step (conversion of HMG-CoA to mevalonate) in cholesterol synthesis. In the drug development process, HMG-CoA reductase was an attractive target in contrast to other late-stage intermediates, not only because of its rate-limiting step but also because inhibition of HMG-CoA reductase does not result in buildup of toxic metabolites. Reduced endogenous cholesterol production results in SREBP activation, which upregulates LDLR gene expression, leading to increased plasma LDL-C uptake, thereby lowering LDL-C plasma levels. The role of statin therapy in lowering plasma levels of LDL-C and the concomitant benefit of reduced CVD risk in a primary and secondary prevention setting, has been examined in multiple trials, some of which were discussed earlier. In 2010, a large meta-analysis of 26 randomized control trials with statin therapy in more than 170,000 subjects showed reduced risk of major coronary events by 22% per mmol/L (40 mg/dl) reduction in LDL-C [61].

TABLE 31.3 Effect of Established Pharmacological Agents on Lipid Fractions

Agent	Lipid Fraction (%)		
	LDL-C	*HDL-C*	*TG*
Statins	18–55↓	5–15↑	7–30↓
Fibrates	5–20↓	10–20↑	20–50↓
Niacin	5–25↓	15–35↑	20–50↓
Bile acid sequestrants	15–30↓	3–5↑	No change
Ezetimibe	18–20↓	1–4↑	8↓

Several statins exist in different classes that vary according to pharmacology and metabolism, as summarized in Table 31.4, as well as in individual effect on the plasma lipid profile. Type 1 statins (lovastatin, simvastatin, and pravastatin) are produced from funghi such as *Aspergillus terreus* by fermentation. Type 2 statins (fluvastatin, atorvastatin and rosuvastatin) are synthetic molecules in which the butyryl group of type 1 statins is replaced with a fluorophenyl group, resulting in increased binding to HMG-CoA. Within this dichotomy, the individual statins differ in their pharmacokinetics depending on absorption/metabolism, lipophilic/hydrophilic rate, and form of administration. With the exception of lovastatin and simvastatin, the statins are administered orally as active hydroxyl acids. Lovastatin and simvastatin are administered as lactone prodrugs that must be hydrolyzed to the active hydroxyl form.

The statins also vary in their hydrophilicity/lipophilicity; pravastatin and rosuvastatin are the most hydrophilic of the statins; the others are lipophilic. Also, with the exception of pravastatin, they undergo phase 1 metabolism in the liver through the CYP450 enzymatic pathways. The particular enzymes responsible are CYP3A (lovastatin, atorvastatin, and simvastatin), CYP2D6 (fluvastatin, pitavastatin, and rosuvastatin), and CYP2C9 (pitavastatin and rosuvastatin). Pravastatin is different in that it is not acted on by the cytochrome P450 system and undergoes metabolism through sulfation and conjugation reactions. The various statin formulations most commonly used are described in Table 31.4 along with their individual pharmacological properties and effect on the plasma lipid profile.

Dosing and Statin Choice

The variety of statins available, and their individual dosing ranges, present a difficulty to prescribers: which statin and dose to choose for which patient? A large meta-analysis recently published compared the six available statins, and found atorvastatin, rosuvastatin, and simvastatin to be the most effective at lowering plasma levels of total cholesterol and LDL-C, with broad equivalence seen at high doses [62]. The study also confirmed the previously known dose–response relationship, whereby a higher dose of the same statin results in a greater reduction of LDL-C [63]. This relationship has been shown to be curvilinear, with a reduction in LDL-C of around 6% seen when the statin dose is doubled. The different statins and their dose-dependent LDL-C-lowering capabilities are detailed in Table 31.5.

Side Effects and Interactions

In general, statin therapy is well tolerated and serious adverse effects have rarely been noted in clinical trials. However, there are known associations with particular side effects, and also interactions with other pharmacological therapy, which prescribers should be aware of when administering statin therapy.

Myopathy

Muscle toxicity (myopathy) related to statin therapy can be a potentially dangerous side effect, as it can progress to rhabdomyolysis and potential renal failure. In the context of statin therapy, myopathy is generally defined as unexplained muscle pain or weakness accompanied by a grossly elevated creatine kinase concentration, >10 times the upper limit of normal. It was first reported in a cardiac transplant patient receiving the immunosuppressive cyclosporine, in addition to gemfibrozil, and, interestingly, was not detected during the original animal safety assessment studies [64].

In August 2001, cerivastatin, which had been introduced in 1998, was withdrawn from the market by its manufacturer

TABLE 31.4 Statins: Pharmacological Properties and Comparative Efficacy

Statin	Metabolism	Hydrophilic	Reduction in TC (%)	Reduction in LDL-C (%)	Increase in HDL-C (%)	Reduction in TG (%)
Lovastatin	CYP3A4	No	16–34	21–42	2–10	6–27
Simvastatin	CYP3A4	No	19–36	26–47	8–16	12–34
Atorvastatin	CYP3A4	No	25–45	26–60	5–13	17–53
Fluvastatin	CYP2C9	No	16–27	22–36	3–11	12–25
Rosuvastatin	CYP2C9	Yes	33–46	45–63	8–14	10–35
Pravastatin	Sulfation	Yes	16–25	22–34	2–12	15–24

TABLE 31.5 Individual Statins and Their Dose-Dependent Effect on Plasma LDL-C Reduction

Statin	Standard Comparator Dose (mg)	Absolute LDL Reduction (95% CI) (mmol/l)	LDL Reduction (%)	Maximum Dose (mg)	Absolute LDL Reduction (95% CI) (mmol/l)	LDL Reduction (%)
Atorvastatin	10	1.79 (1.62–1.97)	37	80	2.64 (2.31–2.96)	55
Fluvastatin	20	1.02 (0.90–1.13)	21	80	1.58 (1.40–1.76)	33
Lovastatin	20	1.40 (1.21–1.59)	29	80	2.15 (1.86–2.43)	45
Pravastatin	20	1.17 (1.10–1.23)	24	40	1.38 (1.31–1.46)	29
Rosuvastatin	20	2.32 (2.20–2.44)	48	40	2.56 (2.42–2.70)	53
Simvastatin	20	1.54 (1.46–1.63)	32	80	2.01 (1.83–2.19)	42

because of a large number of reports of rhabdomyolysis, of which more than 50 cases were fatal [65]. For statins that are substrates for cytochrome P450 CYP3A4 (lovastatin, simvastatin, and atorvastatin), the risk of myopathy is increased by the concomitant administration of potent inhibitors of this enzyme [66].

Myopathy has an incidence of <1 in 1000 patients treated, and a study in 2004 found that in more than 250,000 patients treated with statins there were 24 cases of hospitalized rhabdomyolysis, with similar risks seen between the individual statins and an increased risk associated with concomitant fibrate therapy [60]. The exact mechanism of myopathy has not been fully elucidated. In rat models, it can be prevented by administering mevalonate, the product of the inhibited enzyme [67]. Myopathy risk is higher with advanced age, low body mass, higher doses, concomitant drug administration, and conditions that interfere with hepatic metabolism.

SLCO1B1 and Myopathy

A GWAS of 85 patients with myopathy and 90 matched controls from the 12,000-patient Study of the Effectiveness of Additional Reductions in Cholesterol and Homocysteine (SEARCH) trial identified an association with a common noncoding SNP (rs4149056) in the *SLCO1B1* gene [68]. rs4149056 is in tight LD with a nonsynonymous SNP ($r^2 = 0.97$) that encodes a valine-to-alanine substitution and confers a 4.5-fold increased risk of myopathy per copy of the minor C allele. This variant and others in this highly polymorphic gene have been demonstrated to reduce hepatic uptake and to increase serum statin levels [69]. Roughly 25% of the population carry the C allele, either alone or in combination with other SLCO1B1 variants, all of which appear to similarly affect most statin levels, although the effects are greatest for simvastatin [70]. Myopathy risk is also greatest for simvastatin.

Guidelines for managing simvastatin-induced myopathy risk in the context of SLCO1B1 genotyping were recently published [71]. Although genotyping is not currently required in conjunction with treatment, it is commercially available as a single assay or in combination with other pharmacokinetic variants. Following evidence accumulated on the SLCO1B1 521T>C polymorphism, in 2011 the FDA announced an update to its simvastatin product label recommendations. It is estimated that approximately 30 subjects need to be genotyped to avoid 1 ADR, which has led the Clinical Pharmacogenomics Implementation Consortium (CIPC) to recommend avoidance of the highest simvastatin dose (80 mg) regardless of genotype and use of a dose no higher than 20 mg, along with regular creatine kinase monitoring, in patients known to possess one or more copies of the C allele of rs4149056 (the 521CT or 521 CC genotype).

Hepatotoxicity

Hepatic damage arising from statin therapy has been demonstrated in relevant trials. Hepatotoxicity, defined by an elevation of AST/ALT three times above the upper limit of normal, is seen in around 1–2% of patients on statins, and occurs in a dose-dependent manner. Reversal is seen on dose reduction or cessation of statin therapy.

Diabetes

Results from several studies raised the question of statin therapy potentially having an effect on the incidence of diabetes. The JUPITER trial published results suggesting a causal link between statin therapy and increased incidence of diabetes in the statin arm of the trial, compared to placebo [72]. The opposite finding was reported in an analysis of the WOSCOPS cohort, where statin therapy was associated with reduced frequency of new cases of diabetes in participants receiving pravastatin therapy [73]. A meta-analysis of 13 trials and more than 91,000 participants aimed to resolve this issue, demonstrating a 9% increased risk in incidence of diabetes in patients on statin therapy and a higher risk in older participants [74]. The researchers concluded that this small absolute risk of developing incident diabetes was outweighed by the known cardiovascular benefit conferred by statin use. A prospective cohort study recently published demonstrated a similar risk, which was greater in patients receiving simvastatin, atorvastatin, and rosuvastatin, compared to the reference drug pravastatin [75].

Interactions

Certain drug interactions associated with statin therapy have been identified that increase the likelihood of statin-associated side effects. The majority of these interactions arise through enzyme induction or inhibition of the CYP isoenzymes responsible for statin metabolism, as detailed earlier. It should also be noted that statin therapy could be associated with altered metabolism of concomitant therapies through the same mechanism. A well-established example of this is the elevation of INR seen with statin therapy in patients on warfarin.

Inhibition of the cytochrome P450 isoenzymes can result in increased plasma levels of statins and increased risk of side effects, in particular myopathy and rhabdomyolysis. It is estimated that around a third of patients prescribed a statin are also prescribed a CYP3A4 inhibitor medication, with adverse effects secondary to drug interactions occurring in around 3% of patients [76]. The CYP3A4 isoform, responsible for the metabolism of lovastatin, simvastatin, and atorvastatin, is subject to inhibition by a wide variety of drugs, including macrolide antibiotics, azole antifungals, tricyclic antidepressants, immunosupressants, protease inhibitors, and calcium channel blockers among others. Inducers of the CYP3A4 isoenzyme, such as rifampicin, phenytoin, and St. John's Wort have been identified as reducing the lipid-lowering effect of statins through reduced plasma levels. Similar interactions are seen in statins metabolized by the CYP2C9 isoenzymes.

Outside of the cytochrome P450 system, drug transporters are important mechanisms by which drugs may interact with statin therapy. All statins undergo hepatic uptake through the organic anion-transporting polypeptide (OATP), and several drugs have been identified that either induce or inhibit this pathway. Additional transporters, including P-glycoprotein (P-gp), multidrug resistance-associated proteins (MDR1 and MDR2), breast cancer resistance protein (BCRP), and Na^+-dependent taurocholate cotransporting polypeptide (NTCP), are also subject to similar drug interactions. Several of the statins (atorvastatin, lovastatin, pravastatin, and simvastatin) require glucuronidation through the action of UDP-glucuronosyltransferase 1-1 (UGT-1A) as part of their metabolism, which transforms them into water-soluble products that can be excreted. Inhibitors and inducers of these metabolic pathways, and known statin interactions, are detailed in Table 31.6.

Genetic Variations in the Efficacy of Statins

As mentioned previously, HMG-CoA reductase inhibitors (statins) are associated with a lowering of LDL-C by up to 50%. Despite this, large interindividual variation exists in response to statin therapy. Genetic variations have been identified at various points in the pathways of statin metabolism, affecting the drug's pharmacokinetics and pharmacodynamics. The effect of individual genetic variations on response to statin therapy has been shown to be small and thus the impact on cardiovascular risk has been questioned. A study of the HPS cohort demonstrated an effect size of 0.5–3.0% per allele inherited in 10 common gene variants [77]. In addition, there was no demonstrable effect on cardiovascular risk reduction of statin therapy in association with expression of these common variants.

31.4.3.2 Pharmacokinetic Pharmacogenomics of Statins

As detailed earlier, the enzymes of the cytochrome P450 system are responsible for hepatic phase 1 statin metabolism. Altered pharmacokinetics as a result of genetic variation in the CYP enzymes can lead to altered drug exposure and potentially altered efficacy and increased risk of toxicity.

Various mutations in genes coding for the CYP-metabolizing enzymes have been identified, with varying effects on statin efficacy and toxicity. Polymorphisms in the CYP3A4 gene (rs2740574 and rs4986910), in association with atorvastatin therapy, have been found to have opposite effects on LDL-C [78]. In the case of rs2740574 (M445T variant), carriers of this variant allele at baseline had significantly lower LDL-C levels (−11.2%) relative to noncarriers, and this difference was more pronounced after treatment with atorvastatin (−17.6%). The SNP rs4986910 coded for the A290G variant in a CYP3A4 promotor gene, and homozygotes for this variant were found to have higher LDL-C levels (+12.6%) post-treatment than subjects with 1 or more wild-type allele.

Polymorphisms in the CYP3A5 gene, closely linked to CYP3A4, have been associated with interindividual differences in the efficacy of statin therapy. CYP3A5 is only

TABLE 31.6 Enzyme Inhibitors and Metabolic Pathway Inducers in Statin Metabolism

Metabolic Pathway	Statin	Inhibitors	Inducers
CYP Isoenzymes			
CYP3A4	Atorvastatin, lovastatin, simvastatin	Ketoconazole, itraconazole, fluconazole, macrolides, tricyclic antidepressants, nefazodone, venlafaxine, fluvoxamin, fluoxetine, sertraline, cyclosporin A, tacrolimus, mibefradil, amiodarone, danazol, diltiazem, verapamil, protease inhibitors, midazolam, corticosteroids, grapefruit juice, tamoxifen, sildenafil, warfarin	Phenytoin, phenobarbital, barbiturates, rifampicin, dexamethasone, cyclophosphamide, carbamazepine, omeprazole, St John's Wort
CYP2C9	Fluvastatin, rosuvastatin	Ketoconazole, fluconazole, sulfaphenazole, oxandrolone, amiodarone, dronedarone, warfarin	Rifampicin, phenobarbital, phenytoin
Transporter Proteins			
MDR/P-gp	Atorvastatin, lovastatin, pravastatin, simvastatin	Ritonavir, ciclosporin, verapamil, erythromycin, ketoconazole, itraconazole, quinidine, elacridar	Rifampicin, St John's Wort
OATP1B1	All statins	Ciclosporin, rifampicin, gemfibrozil, macrolides, indinavir, ritonavir, saquinavir	
Uridine Diphosphate Glucuronyltransferase			
UGT-1A	Atorvastatin, lovastatin, pravastatin, simvastatin	Gemfibrozil, ciclosporin	Rifampicin

CYP—cytochrome P450; MDR—multidrug resistance-associated protein; P-gp—P-glycoprotein; OATP—organic anion–transporting polypeptide; UGT—Uridine 5′-diphospho-glucuronosyltransferase.

expressed in around 10% of Caucasian subjects, and it was postulated that individuals carrying at least one copy of the wild-type CYP3A5*1 allele would show increased statin metabolism, and therefore diminished lowering of LDL-C, relative to those homozygous for the CYP3A5*3 allele deemed nonexpressors of CYP3A5 [79]. In subjects expressing CYP3A5*1, LDL-C levels were higher by 24% than in those CYP3A5*3 homozygotes after one year on treatment with lovastatin, simvastatin, or atorvastatin. It was concluded that patients expressing the CYP3A5*1 allele may require increased doses of statins to achieve the desired hypolipidemic effect.

Genetic variation in CYP2D6 has been associated with variance in both statin efficacy and tolerance. Patients with polymorphisms in CYP2D6, resulting in a status of "poor metabolizers," have been shown to have enhanced lowering of total cholesterol with simvastatin therapy [80]. Unfortunately, decreased simvastatin metabolism resulting in increased drug efficacy also had the effect of increased toxicity, with more patients discontinuing treatment. In addition, a particular polymorphism, the CYP2D6*4 allele (rs3892097), was identified in a study on atorvastatin-induced myopathy as being present in 50% of those affected versus only 28% of controls [81].

CYP7A1 encodes for the 7α-hydroxylase enzyme, which is the catalyst for the rate-limiting step in the synthesis of bile acid from cholesterol. Carriers of the A-204C promoter variant have been shown to experience a significantly reduced efficacy of pravastatin therapy, with an LDL-C lowering effect of 24.3% versus 33.1% in noncarriers [82].

The polymorphisms just discussed are found in the proteins responsible for statin metabolism. The pharmacokinetics of statin therapy is also dictated by the uptake of the drug in a process governed by the membrane uptake transport proteins, in particular the organic anion-transporting polypeptides (OATPs) or the solute carrier organic anion transporters (SLCOs). The focus for much of the investigation surrounding these proteins is centered on the *SLCO1B1* gene, which codes for the OATP1B1 influx transporter and its role in statin-induced myopathy (see earlier).

31.4.3.3 Pharmacodynamic Pharmacogenomics of Statins

The genetic variations detailed earlier concern factors that dictate the pharmacokinetic properties of statins. This section examines variations that govern the pharmacodynamic properties of statins and, in turn, affect the extent to which statins reduce cardiovascular risk. (See Figure 31.4.)

The first step in cholesterol synthesis, and the pharmacological target for statins, is HMG-CoA reductase. It therefore follows that genetic variation in HMG-CoAR potentially leads to variability in statin therapy response. In a study of 1536 patients treated with pravastatin in the PRINCE cohort, two SNPs, 12 and 29 (rs17244841 and rs17238540), were identified with reduced response [83]. These SNPs are in tight linkage disequilibrium with each other. Subjects with one copy of the minor allele of either were found to have a 22% lower reduction in total cholesterol and 19% lower reduction in LDL-C relative to subjects homozygous for the major allele for one of them. A further study of rs17238540 confirmed reduced statin therapy efficacy in subjects expressing the less frequent G allele [84].

The ATP-binding cassette (ABC) transporters G5 and G8 play a role in the intestinal absorption, hepatic transportation, and excretion of cholesterol. The D19H variant (rs11887534) in the *ABCG5/8* gene has been shown to be associated with a greater LDL-C-lowering response with atorvastatin, than the wild-type variant [85].

Apolipoprotein E (Apo E) is present in triglyceride-rich lipoproteins and HDL, and interacts with the LDL receptor to promote lipoprotein clearance. The Apo E gene is polymorphic with three alleles, E2, E3, and E4, in order of increasing affinity for LDL-R. Various studies have shown how variations in Apo E alleles alter lipid levels and potentially CVD risk, as described earlier. Several studies have shown variance in response to statin therapy depending on the Apo E allele expressed, with the E2 allele being associated with an enhanced response and, conversely, the E4 allele being associated with an attenuated hypolipidemic response [86]. It should be noted, however, that results on this relationship are inconsistent and that certain studies have shown that the Apo E genotype loses significance for variable response to therapy when other factors such as BMI, age, and sex are adjusted for. The effect of Apo E polymorphisms and statin therapy on cardiovascular outcomes has also been investigated.

In an analysis of the 4S cohort, the E4 allele was associated with a nearly two-fold increase in mortality relative to noncarriers [87]. This excess risk was found to be almost entirely abolished by statin treatment. However, this finding was not replicated in a further study of the Regression Growth Evaluation Statin Study (REGRESS) cohort, where there were no improvements in angiographic parameters of CAD associated with the Apo E genotype [88]. It can be concluded that variants in the Apo E genotype are associated with varying response to statin therapy, but that the evidence for any effect on cardiovascular outcomes is equivocal.

The role of the cholesterol ester transfer protein (CETP) in lipid metabolism is described in detail elsewhere in this chapter. The most investigated CETP variant is the SNP rs708272, also known as the Taq1B variant. The B1 allele is associated with elevated CETP levels, resulting in lower plasma HDL-C levels, and subjects homozygous for the B1 allele have accelerated atherosclerosis compared to wild-type carriers. Interestingly, studies have shown that subjects with the B1B1 genotype have a more atherogenic lipid profile at baseline, but they respond well to statin therapy and thus incur lower mortality from cardiovascular events than subjects with the seemingly less atherogenic B2 genotype [89].

Statin therapy has an indirect effect, through proprotein convertase subtilisin/kexin type 9 (PCSK9), on LDL-C levels other than that mediated by HMG-CoAR, and polymorphisms in the PCSK9 gene may alter response to statin therapy. Statin-induced upregulation of the sterol regulatory element-binding protein-2 (SREBP-2) results in activation of the *PCSK9* gene, leading to increased plasma PCSK9 levels. It follows therefore that polymorphisms in the *PCSK9* gene result in varying statin response, depending on whether they are "loss-of-function" or "gain-of-function." The clinical consequences of this are described in a following section on PSCK9 as a pharmacological target.

Kinesin-like protein 6 (KIF6) is responsible for the intracellular transport of several molecules, including mRNA. Initial studies in large clinical-trial cohorts identified an association between the rs20455 SNP (Trp719 Arg), increased risk of CVD and increased statin efficacy [90]. In the WOSCOPS cohort, carriers of the 719Arg allele had an absolute CAD risk reduction of 5.5% from statin therapy, with noncarriers having a risk reduction of 0.1%. This benefit of the 719Arg polymorphism was echoed in an analysis of the PROSPER, CARE, and PROVE IT-TIMI 22 cohorts. However, these positive findings have not been replicated in further studies, and a large meta-analysis of 17,000 cases of CAD showed no association between the 719Arg allele and CAD risk compared to noncarriers [91]. This was also the case in the HPS study cohort, where no association between KIF6 polymorphisms and CVD risk or statin response was elicited [92].

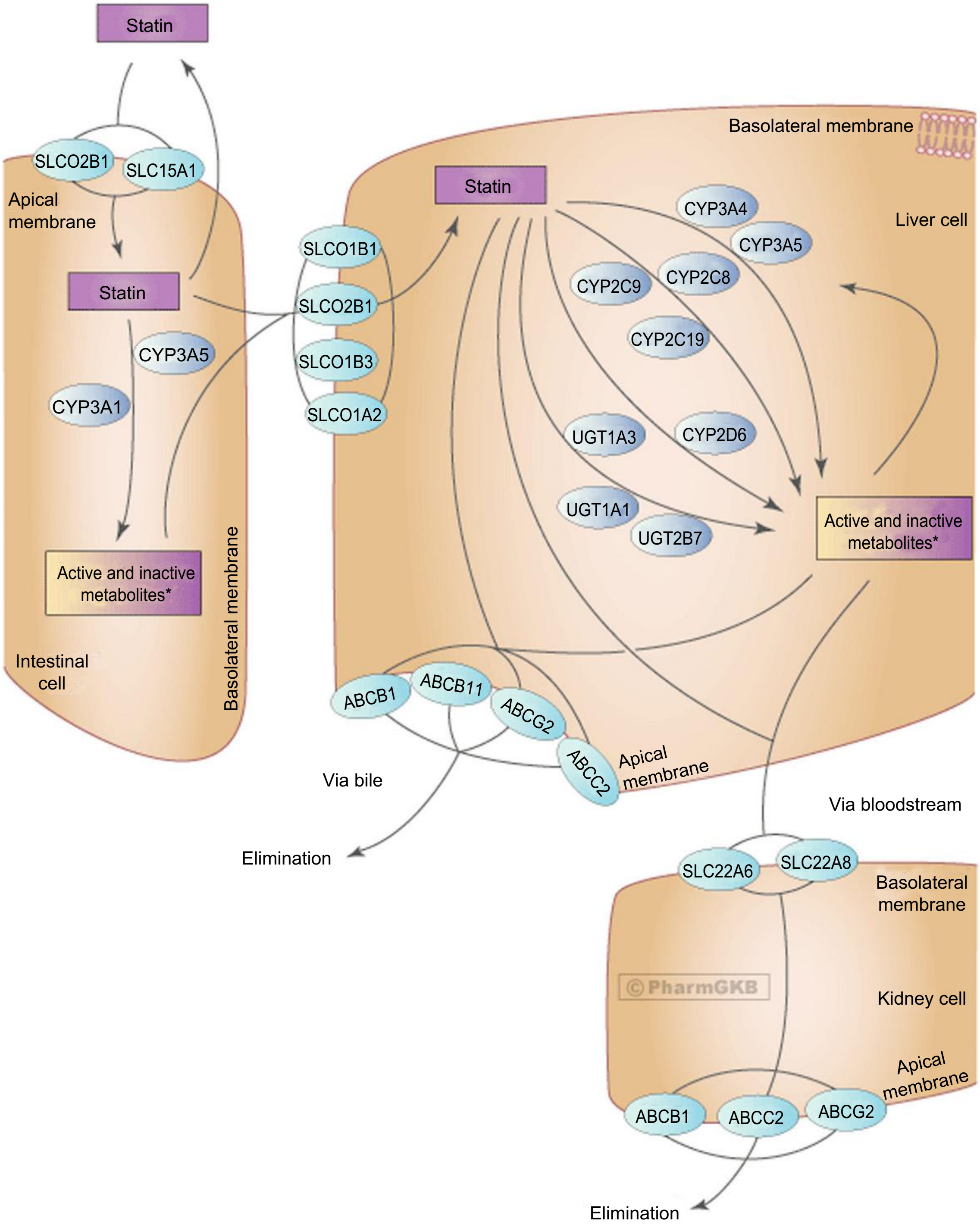

Statin
SLCO2B1
SLC15A1
Apical membrane
Statin
CYP3A5
CYP3A1
Active and inactive metabolites*
Intestinal cell
Basolateral membrane
SLCO1B1
SLCO2B1
SLCO1B3
SLCO1A2
Basolateral membrane
Statin
CYP3A4
CYP3A5
Liver cell
CYP2C9
CYP2C8
CYP2C19
UGT1A3
CYP2D6
UGT1A1
UGT2B7
Active and inactive metabolites*
ABCB1
ABCB11
ABCG2
ABCC2
Apical membrane
Via bile
Elimination
Via bloodstream
SLC22A6
SLC22A8
Basolateral membrane
© PharmGKB
Kidney cell
Apical membrane
ABCB1
ABCC2
ABCG2
Elimination

FIGURE 31.4 **Statin pharmacokinetics.** Statins are oral inhibitors of HMG-CoA reductase. They enter the systemic circulation by intestinal absorption and transportation through the action of the ABC and SLC gene transporters. Metabolism is mainly hepatic and is driven by the CYP and UGT gene enzymes. Excretion occurs through the biliary route, mediated by the ABC transporters. *Source: Reproduced with permission from PharmGKB and Stanford University (copyright held by PharmGKB).*

Ezetimibe

Ezetimibe is a selective inhibitor of cholesterol absorption that has been shown to be an effective LDL-C lowering agent as monotherapy or in combination with statin therapy.

Ezetimibe is taken as a once-daily oral medication and binds to the intestinal cholesterol transporter Niemann-Pick C1-Like 1 (NPC1L1). The binding of ezetimibe to the transporter protein prevents the vesicular internalization of NPC1L1 and cholesterol into the small intestine, thereby inhibiting its absorption (Figure 31.5). In addition, decreased cholesterol absorption leads to upregulation of genes involved in LDL-C uptake and de novo sterol synthesis, in particular SREBP-2, LDL-R, and HMG-COAR, leading to increased cellular clearance of LDL-C from the plasma.

In monotherapy, ezetimibe was calculated to lower plasma levels of LDL-C, with a mean reduction of 18.58% compared to placebo, in a meta-analysis of eight trials

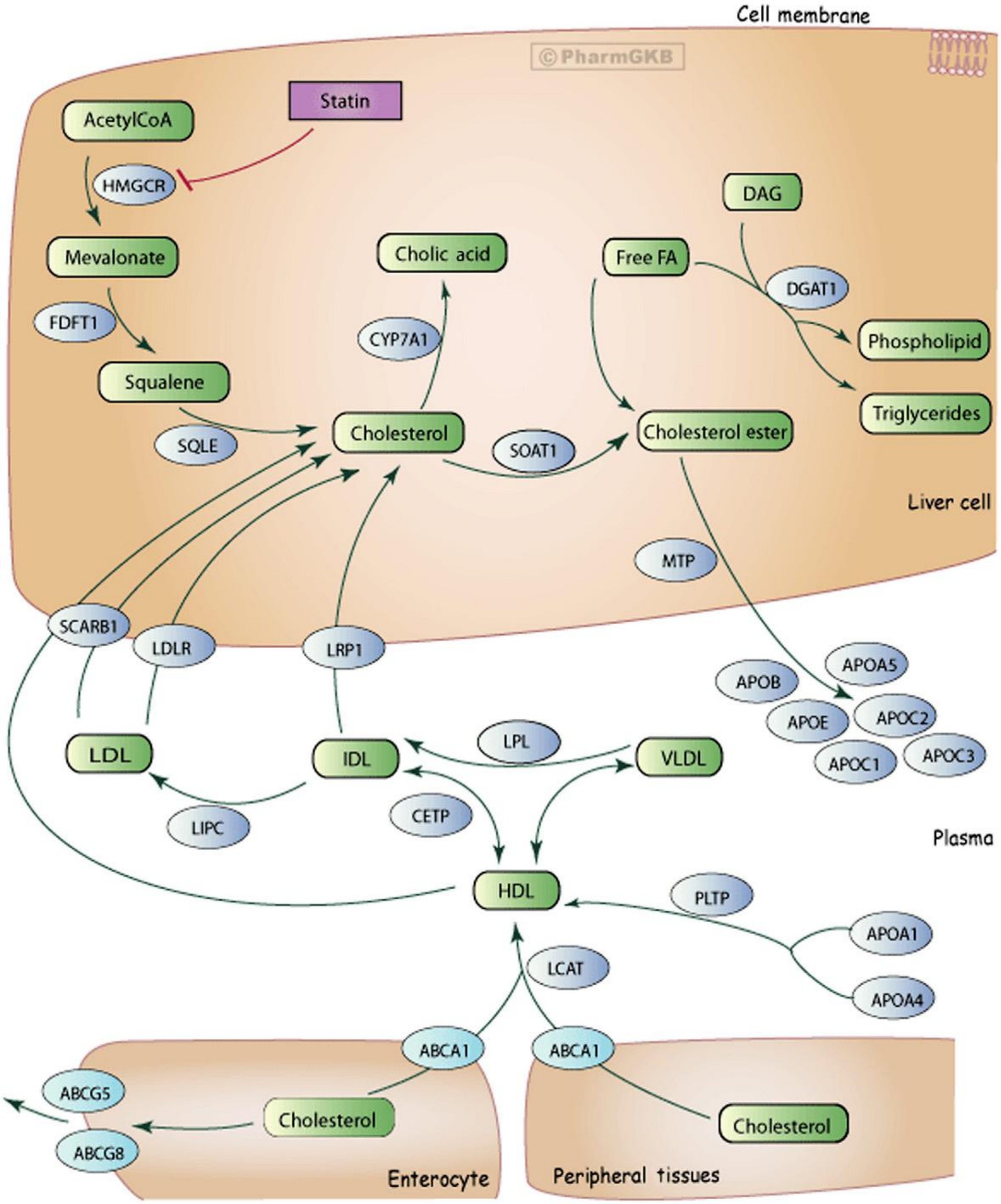

Cell membrane
© PharmGKB
Statin
AcetylCoA
HMGCR
Mevalonate
FDFT1
Squalene
SQLE
Cholic acid
CYP7A1
Cholesterol
SOAT1
DAG
Free FA
DGAT1
Phospholipid
Triglycerides
Cholesterol ester
Liver cell
MTP
SCARB1
LDLR
LRP1
APOA5
APOB
APOE
APOC2
APOC1
APOC3
LDL
IDL
LPL
VLDL
LIPC
CETP
Plasma
HDL
PLTP
APOA1
APOA4
LCAT
ABCA1
ABCA1
ABCG5
ABCG8
Cholesterol
Cholesterol
Enterocyte
Peripheral tissues

FIGURE 31.5 **Statin pharmacodynamics.** Statins affect plasma LDL-C levels through inhibition of HMG-CoA reductase, resulting in reduced endogenous production of cholesterol. Various genes affect the pharmacodynamics of statin therapy, and variations in these gene products can alter therapy efficacy and safety. *Source: Reproduced with permission from PharmGKB and Stanford University (copyright held by PharmGKB).*

involving 2722 people [93]. Several studies demonstrated a benefit in combination therapy of ezetimibe with a statin [94,95]. In patients with primary hypercholesterolemia, the addition of ezetimibe proffered an additional 12% reduction in LDL-C compared to monotherapy with atorvastatin, and this finding was echoed in a combination of ezetimibe with simvastatin, where an additional reduction in LDL-C of 13.8% was seen. In the setting of familial hypercholesterolemia, ezetimibe's effect on a surrogate end point of atherosclerosis, carotid-intima thickness (cIMT), was examined with unexpected findings [96]. Despite significant reductions in LDL-C levels, ezetimibe combination therapy with 80 mg of simvastatin was associated with an insignificant increase in cIMT, in comparison to simvastatin monotherapy. In terms of reduction of cardiovascular events, the Study of Heart and Renal Protection (SHARP) demonstrated a 17% reduction with ezetimibe plus simvastatin in the rate of major atherosclerotic events in comparison to placebo [97].

It should be noted that this study was carried out in a select group of patients who had advanced chronic kidney disease. The large Improved Reduction of Outcomes: Vytorin Efficacy International Trial (IMPROVE-IT) recruited 18,000 patients to examine the combination of ezetimibe and simvastatin in patients after acute coronary syndrome. The results are due to be released in 2014 and are hoped to give important information on the effect and safety profile of LDL-C lowering with combination therapy. Concerns were raised about the safety of ezetimibe after the Simvastatin and Ezetimibe in Aortic Stenosis (SEAS) trial found a significantly higher onset of cancer in its treatment arm. This finding, however, was refuted in analysis of the SHARP cohort and also of the data available from IMPROVE-IT [98].

In conclusion, ezetimibe is effective at lowering plasma levels of LDL-C in monotherapy, or in combination with statins. It has been shown to result in lower risk of cardiovascular events, albeit in a select group of patients, and further trial data are awaited with regard to its effect on cardiovascular end points. Current guidelines endorse ezetimibe in addition to statin therapy when LDL-C targets are not met with the maximal tolerated statin dose and when patients are intolerant of statins or have contraindications to them [24].

As mentioned previously, the target for ezetimibe is the cholesterol transporter NPC1L1. Genetic variants in the gene coding for this protein have been identified, with resulting interindividual variety in response to ezetimibe. The possibility of genetic variance in NPC1L1 was first raised when the genomic data of eight individuals not responsive to ezetimibe in terms of LDL-C reductions were examined [99]. One individual possessed two novel polymorphisms in the NPC1L1 gene, V55L and I1233N, both of which were absent in the control group. The authors proposed that compound heterozygosity in this subject was the reason for his lack of response.

In another study, 101 subjects with dyslipidemia were examined with regard to their NPC1L1 genotype and treated with ezetimibe for a total of 12 weeks [100]. Subjects lacking the haplotype 1735C-25342A-27677T were found to have a significantly greater reduction in LDL-C versus those who had at least one copy of the haplotype (35.9% versus 23.6%; $p = 0.0054$). The conclusion reached was that genetic variants in NPC1L1 are responsible for interindividual variability in response to LDL-C lowering by ezetimibe.

Bile Acid Sequestrants

As detailed earlier in the chapter, bile acids are produced in the liver from the cytochrome P450-mediated oxidation of cholesterol. They are then secreted into the intestinal lumen to be either excreted or reabsorbed into the enterohepatic circulation. Bile acid sequestrants (BASs) are one of the oldest lipid-lowering medications available. They act as positively charged nonabsorbable resins that can bind to the negatively charged bile salts, preventing reabsorption into the enterohepatic circulation. Reduced reabsorption of bile acids results in upregulation of hepatic CYP7A1, causing increased oxidation of cholesterol into bile salts, and a depletion of intrahepatic cholesterol. Reduced intracellular cholesterol results in SREBP-driven increased expression of HMG-CoAR and LDLR, with a net effect of reduction in plasma LDL-C concentrations.

Three BASs are available: cholestyramine, colestipol, and colesevelam-HCL. Colesevelam differs slightly from the other first-generation BASs in that it has been synthetically engineered to contain longer hydrophobic side chains, resulting in its increased binding of bile salts. In addition to lowering LDL-C, colesevelam has been shown to reduce HbA1c in patients with type 2 diabetes mellitus [101]. BASs are unfortunately associated with a common frequency of side effects even at low dose, which limits their use in clinical practice. These include gastrointestinal upset such as flatulence, constipation, dyspepsia, and nausea. Another adverse effect of BAS therapy is the reduced absorption of several drugs and fat-soluble vitamins (A, D, E, and K), which has led to the advice that any BAS should be taken four hours before other drugs or one hour after.

The Lipid Research Clinics Coronary Primary Prevention Trial showed a reduction in LDL-C of 20% with cholestyramine therapy versus placebo, with a resultant significant 19% reduction in fatal and nonfatal MIs [102]. This finding has been echoed in further clinical trials for all BASs, with a reduction in LDL-C of 18–25% at maximum dose. The addition of BASs to statin therapy results in an additional LDL-C-lowering effect of 10–15%. The combination of ezetimibe, a statin, and colesevelam was shown to result in significantly increased LDL-C reduction when compared to combination therapy with statin and ezetimibe alone, in a cohort of patients with familial hypercholesterolemia and baseline LDL-C levels >2 mmol/L [103].

Bile acid sequestrants are currently recommended as second-line agents in addition to statin therapy, where targets are not met with maximum tolerated dose of statins or when the patient cannot tolerate them [24].

Novel Medications

Despite statins' universally accepted LDL-C-lowering effect, it is thought that up to 30–50% of patients with moderate to high risk of CVD do not achieve target goals for non-HDL-C [104]. Work is ongoing to discover novel medications that help to further lower LDL-C in patients with very high CVD risk (e.g., familial hypercholesterolemia) or that help patients who are statin-intolerant.

Microsomal Triglyceride Transfer Protein Inhibitors

Microsomal triglyceride transfer protein (MTP) is responsible for the hepatic assembly of VLDL particles through the transfer of triglycerides to apolipoprotein B. Mutations in the MTP gene are known to lead to a deficiency in apolipoproteins B-48 and B-100, known as abetalipoproteinemia, in which sufferers are hypocholesterolemic. It is thought that pharmacological inhibition of MTP results in hepatic production of VLDL, the precursor to LDL, causing a lowering of plasma LDL-C levels.

Lomitapide (BMS-201038) is the only MTP inhibitor to have been investigated in clinical trials. In a small trial of six patients with homozygous familial hypercholesterolemia, lomitapide therapy resulted in an LDL-C reduction of 50.9% [105]. However, one-third of the cohort experienced derangement in liver function tests. This is thought to be due to a side effect of MTP inhibition that causes increased hepatic fat stores and results in a state analogous to nonalcoholic fatty liver disease.

To examine the safety of MTP inhibition further, investigators carried out another study with lomitapide in 29 patients with FH [106]. A similar reduction in LDL-C to that in the first study was found, with 93% of patients reporting gastrointestinal upset. Derangement of liver transaminases greater or equal to three times the upper limit of normal were seen in ten patients, with four demonstrating derangement 5 or more times greater than the upper limit of normal. All derangements were managed with dose reduction or temporary discontinuation, with no permanent discontinuation necessary. The results of this trial led the U.S. Food and Drug Administration (FDA) to approve lomitapide as a therapy for patients with homozygous familial hypercholesterolemia. No surrogate or cardiovascular end-point data for lomitapide is yet available.

Antisense Oligonucleotide Therapies

Mipomersen is an antisense oligonucleotide therapy that specifically binds to the apolipoprotein B-100 mRNA, blocking translation of the gene product. The reduction in production of apo B-100 results in reduced hepatic production of the atherogenic lipoproteins VLDL, IDL, LDL, and lipoprotein(a). Mipomersen has been shown to reduce LDL-C in addition to standard therapy in patient cohorts with homozygous/heterozygous FH, in those with preexisting CAD, and in those intolerant of statin therapy and at high risk of CVD [107,108]. Mipomersen therapy was associated with an additional reduction in LDL-C of 25–37% from baseline. In addition, mipomersen therapy resulted in reduction in levels of apo-B, Lp(A), non-HDL-C, and triglycerides. Some concerns have arisen regarding the effect of mipomersen on liver architecture, with persistent rises in transaminases showing evidence of hepatic steatosis on MRI. Mipomersen is approved by the FDA for therapy in patients with homozygous and severe heterozygous familial hypercholesterolemia.

PCSK9

The effect of mutations in proprotein convertase subtilisin kexin type 9 (PCSK9) on lipid homeostasis were first described in 2003 [31]. Over the following years, the exact mechanism of PCSK9 was elucidated; several more mutations, and their effect on LDL-C levels, have been described. PCSK9 is a serine protease that is secreted by the liver. It acts to regulate plasma LDL-C levels through post-translational downregulation of the LDL-receptor (LDLR), by increased lysosomal degradation. On secretion, the prodomain is autocatalytically cleaved and continues to remain associated with the catalytic domain of the protein; it is thought the reduction of LDLR by PCSK9 is independent of any enzymatic activity.

The role of PCSK9 in cholesterol homeostasis was determined through the discovery of two mutations in the *PCSK9* gene, which led to a third locus associated with autosomal dominant hypercholesterolemia in addition to LDLR and APOB. Shortly after the publication of these findings, it was found that overexpression of PSCK9 in mice, using an adenovirus constitutively expressing murine PCSK9, resulted in a two-fold increase in plasma total cholesterol, a five-fold increase in non-HDL cholesterol, and a near complete absence of LDLR. The reduction in LDLR appeared to be independent of any effect on LDLR mRNA levels, confirming that the effect of PCSK9 on LDLR levels was at a post-translational level [109].

Mutations in PCSK9, and the consequent increase in plasma LDL-C levels and so increased CVD risk, were recognized to be due to the gain-of-function mutations. The potential role of PCSK9 as a therapeutic target arose following the publication by Cohen et al. in 2006 of the discovery of loss-of-function mutations in PCSK9 and an associated reduction in cardiovascular risk through low levels of LDL-C [35]. Cohen et al. reported that approximately 2% of African-Americans were heterozygous for one of two loss-of-function nonsense mutations (Y142X or C679X) in PCSK9, with a resultant 40% reduction in mean LDL cholesterol. This reduction proffered an 80–90% reduction in the incidence of coronary heart disease. Similar results were found in Caucasian subjects through a novel mutation, R46L, resulting in a 15% reduction in LDL-C and a 50% reduction in risk of coronary heart disease.

To date, two methods of PCSK9 inhibition have been investigated. The first is the use of an antisense oligonucleotide (ASO) targeting PCSK9 mRNA. Such a method has been shown to reduce levels of total cholesterol in both mice and nonhuman primates by up to 50% [110], achieved through a lowering of plasma levels of PCSK9 mRNA and protein with increased hepatic expression of LDL-R protein. The second method has been to target the binding of PCSK9 and LDLR through use of humanized anti-PCSK9 monoclonal antibodies. Several such drugs are in ongoing clinical trials, and early results appear promising. In comparison to placebo, injection of the monoclonal antibody REGN727 in patients with familial hypercholesterolemia resulted in marked reductions in LDL-C with no significant adverse effects [111].

In another study, the combination of REGN727 with 80 mg of atorvastatin was associated with a 73.2% reduction in LDL-C, in comparison to 17.3% in those receiving atorvastatin 80 mg plus placebo [112]. A similar reduction (66.2%) was seen in those receiving a lower dose of atorvastatin (10 mg) in addition to REGN727. Subjects in the study treatment group also demonstrated significant reductions in Apo B, non-high-density lipoprotein-C, lipoprotein(a), and triglycerides, with no significant adverse findings, although further, adequately powered studies are required to fully assess the safety profile of such drugs.

The studies just described show promise in the therapeutic lowering of PCSK9 levels and subsequent reduction in LDL-C levels greater than that currently achieved with statin therapy alone. In years to come, this approach will potentially play a vital role in cholesterol-lowering therapy in patients intolerant of statins and in those for whom statin therapy does not achieve target levels of LDL-C, with the aim of reducing cardiovascular risk beyond that currently possible.

31.4.4 HDL-C

The role of elevated LDL-C and low levels of plasma HDL-C in increasing cardiovascular risk were described in detail earlier. The reduction in LDL-C levels through statin therapy has been shown to significantly reduce the incidence of and mortality associated with coronary heart disease [10]. However, many patients still show increased residual risk along with increased incidence of cardiovascular events despite adequate statin therapy, particularly with the rising prevalence of obesity, type 2 diabetes mellitus, and the metabolic syndrome.

Several studies have demonstrated that low HDL-C levels are an independent risk factor for cardiovascular events [13,14]. Elevating HDL-C levels to reduce cardiovascular risk, in addition to reductions gained by lowering LDL-C, has been an important target for pharmacological therapy as an adjunct to statins. To date, this potential avenue for lowering CVD risk has not been successful, leading to questions concerning the role of HDL-C as a causal factor in CVD risk [56].

31.4.4.1 CETP

Similar to the development of PCSK9 inhibitors, the Mendelian disorder of CETP deficiency was the stimulus for the development of cholesteryl ester transfer protein (CETP) inhibitors. CETP plays a vital role in the atheroprotective process of reverse cholesterol transportation, which was described in detail earlier. In summary, CETP involves the HDL-mediated transport of cholesterol from the peripheral tissues to the liver for recycling or excretion as bile acids. It is secreted by hepatic and adipose tissue and is circulated in the plasma, bound to HDL-C molecules. It aids the transfer of cholesterol esters, in exchange for triglyceride molecules, from HDL to VLDL and LDL. Patients with CETP deficiency have markedly elevated HDL-C levels, and it was thought that pharmacological inhibition of CETP could induce a state analogous to CETP deficiency, with the aim of increasing HDL-C and potentially reducing CVD risk. Observational data from two large genome analyses showed an association of SNPs around the CETP locus with HDL-C levels and CVD risk, with inhibition of CETP resulting in higher HDL-C levels and a concordant CVD risk reduction [113,114].

To date, several CETP inhibitors have been investigated with varying results. The first, torcetrapib, was an irreversible CETP inhibitor that was found to increase HDL-C levels by up to 106% at maximum dosage [115]. In an effort to ascertain the clinical effect of torcetrapib, the trial known as Investigation of Lipid Level Management to Understand its Impact in Atherosclerotic Events (ILLUMINATE) randomized 15,067 subjects with a background of CVD or type 2 diabetes to torcetrapib plus atorvastatin or placebo plus atorvastatin [116]. The trial was stopped only 1.5 years into the planned 4.5-year follow-up as there was found to be a significantly increased rate of cardiovascular events in the torcetrapib group versus placebo (464 versus 373; HR 1.25; 95%CI 1.09–1.44). This was despite a seemingly beneficial change in lipid profile, with an increase in HDL-C of 72.1% and a decrease in LDL-C of 24.9% from baseline at 12 months. It was additionally found that systolic blood pressure in the torcetrapib group was significantly raised by 5.4 mmHg versus placebo, and that there were significant decreases in serum potassium and increases in serum bicarbonate, sodium, and aldosterone. It was thought that the excess mortality and changes in blood pressure found were possibly due to off-target effects of torcetrapib on the renin–angiotensin–aldosterone system.

Dalcetrapib is a weaker CETP inhibitor that was found to increase HDL-C by around 30% with very little effect on LDL-C [117]. Initial phase II safety trials were promising, with dalcetrapib therapy being well tolerated without significant changes in blood pressure or any change in atherosclerotic plaque progression/vascular inflammation as visualized by PET-CT imaging [118]. The dal-OUTCOMES trial was a large multicenter placebo-randomized control trial involving 15,871 patients with a recent acute coronary syndrome [119]. These patients were randomized to dalcetrapib or placebo in addition to standard evidence-based statin therapy. Dalcetrapib was associated with a significant increase in HDL-C levels compared to placebo; however, the trial was stopped at interim analysis on the grounds of futility given that treatment with dalcetrapib had no significant effect on cardiovascular outcomes. Interestingly, the researchers found no association between baseline HDL-C levels and risk of cardiovascular events in either the treatment or the placebo arm of the trial cohort. The trial did not find any concerns with regard to adverse effects associated with dalcetrapib, except for a small (0.6 mmHg) increase in systolic blood pressure in comparison to placebo.

The third CETP inhibitor to be investigated, anacetrapib, is a powerful CETP inhibitor that raises HDL-C by 138% and lowers LDL-C by 40%, compared to placebo, with no apparent safety concerns [120]. Its role as an adjunct to optimal lipid-lowering therapy on cardiovascular outcomes is currently being investigated in the Randomized Evaluation of the Effects of Anacetrapib through Lipid-modification (REVEAL) study, the results of which are expected in 2017. Evacetrapib, the fourth CETP inhibitor, is similarly being investigated in a large phase III outcome trial in the Assessment of Clinical Effects of Cholesteryl Ester Transfer Protein Inhibition with Evacetrapib in Patients at a High-Risk for Vascular Outcomes (ACCELERATE) trial. Phase II studies found evacetrapib to have an effect on the lipid profile, with a rise in HDL-C of 79–89% and a decrease in LDL-C of 11–14% in conjunction with statin therapy [121].

To date, the cardiovascular outcome data on CETP inhibitors have been disappointing, with two large phase II trials failing to reach completion because of adverse effects (torcetrapib) and futility (dalcetrapib). It is hoped that the two trials currently under way will show the benefit of this novel class of drugs on cardiovascular outcomes through modification of the plasma lipid profile, in particular its action on HDL-C.

31.4.4.2 Nicotinic Acid (Niacin)

Nicotinic acid, better known as Niacin, was first noted, in the 1950s, to cause decreased levels of total cholesterol and LDL-C along with increased HDL-C. Later, it was also shown to cause decreased levels of triglycerides and lipoprotein(a). Niacin's role as a therapeutic intervention in the management of lipid abnormalities, in particular raising HDL-C levels, has been extensively investigated with varying results.

Niacin alters plasma lipid parameters through various mechanisms. It exerts an anti-lipolytic action through activation of the nicotinic acid receptor HCA_2 (also known as G protein-coupled receptor (GPR109A), resulting in inhibition of cyclic adenosine monophosate (cAMP) and a decreased release of free fatty acids from adipose tissue. This causes decreased hepatic synthesis of triglycerides, along with VLDL particles, and consequently decreased levels of LDL-C. Additionally, the reduction in plasma supply of free fatty acids to the hepatic tissue has been shown to result in decreased hepatic expression of peroxisome proliferator-activated receptor γ (PPARγ), coactivator-1β (PGC-1β), and apolipoprotein C-3 (APOC3). This has the consequence of reduced hepatic production and increased clearance of VLDL particles.

Niacin has also been shown to decrease hepatic triglyceride synthesis directly, through inhibition of diacylglycerol acyltransferase-2. Its effect on the plasma HDL-C profile is less clear but is thought to have several aspects. It has been shown to cause decreased hepatic clearance of ApoA1, the major component of HDL-C, resulting in increased reentry of HDL-C into the circulation from hepatic cells. It has also been shown to increase plasma HDL-C through decreased CETP gene expression, resulting in decreased exchange of cholesterol esters (HDL-C) and triglycerides (LDL-C and VLD-L). Niacin has also been purported to have effects on the atherosclerotic process, independent of lipid profile modification. It has been shown to inhibit expression of vascular cell adhesion molecule 1 (VCAM1), resulting in stabilization of atherosclerotic plaques alongside an anti-thrombotic effect. In addition, niacin-mediated activation of macrophages expressing the HCA_2 receptor results in increased transcription and activity of PPARγ, and the cholesterol transporters ABCA1 and ABCG1. This process increases the efflux of cholesterol from atherosclerotic plaques to HDL-C through reverse cholesterol transport.

As detailed in the preceding paragraph, the receptor GPR109A is responsible for the lipid-lowering mechanism of niacin. GPR109A is coded for by the HM74A gene and a gene adjacent to it, HM74, has been found to code for a similar receptor, GPR109B. Genomic analysis has demonstrated variance in the haplotypes for these two genes. However, to date there is no data suggesting any interindividual variability with regard to response to niacin [122].

The role of niacin in a clinical setting with the purpose of reducing cardiovascular end points by its modification of the plasma lipid profile has been intensively studied. The hope is that it can help to reduce the residual risk that remains despite intensive lowering of LDL-C by statin therapy. Niacin has an effect on the whole spectrum of dyslipidemia, with a rise in HDL-C of 15–35% seen, along with a reduction in LDL-C and triglycerides by 15–18% and 20–40%, respectively. One of the first studies to show a potential benefit of niacin therapy was the Coronary Drug Project (CDP), which found that, in a population of 8341 males after myocardial infarction, niacin therapy resulted in no reduction in overall mortality versus placebo. However, it did display a statistically significant 14% relative risk reduction in nonfatal MI versus placebo.

A 15-year follow-up study of the niacin treatment arm showed an 11% reduction in overall mortality [123]. Several studies have examined the effect of niacin on surrogate end points of atherosclerosis, in particular carotid intima-media thickness (cIMT)—a measurement of atherosclerotic plaque thickness using ultrasound. Arterial Biology for the Investigation of Treatment Effects of Reducing Cholesterol (ARBITER) 3, the third in a series of trials examining the effect of niacin in addition to simvastatin therapy on cIMT [124]—demonstrated a significant reduction in cIMT in subjects receiving niacin and simvastatin versus placebo and simvastatin. Following these trials, which suggested a potential positive benefit on cardiovascular outcomes and surrogate markers from niacin therapy, two large clinical trials were designed with the aim of proving a benefit of niacin in combination with statin therapy.

AIM-HIGH was a double-blind randomized control trial (RCT) that examined the effect of simvastatin (± ezetimibe to achieve LDL-C <2.07 mmol/l) plus niacin, versus simvastatin (± ezetimibe) plus placebo, on the primary composite end point of death from coronary heart disease, nonfatal myocardial infarction, ischemic stroke, hospitalization for an acute coronary syndrome, or symptom-driven coronary or cerebral revascularization [125]. The trial was stopped early because of a lack of proven efficacy of niacin versus placebo after 36 months of follow-up, despite significant rises in HDL-C in the niacin group versus placebo. A much larger trial, HPS2-THRIVE, looked at 25,673 patients receiving niacin plus the antiflushing agent laropiprant in addition to statin therapy. It failed to show a benefit of niacin therapy on the primary outcome of a combination of coronary death, nonfatal myocardial infarction, stroke, or coronary vascularization. In addition, an increase in adverse events was seen in the treatment arm, with a significant increase in diabetic complications and new-onset diabetes.

Unfortunately, niacin therapy is associated with a high incidence of the side effect of flushing due to the activation of the nicotinic acid receptor HM74A and the subsequent release of prostaglandins PGI_2, PGD_2, and PGE_2. Interestingly, in a study of mice lacking the gene (PUMA-G) coding for GPR109A, it was found that the mice lacked any flushing response to niacin [126]. This suggests that any genetic predisposition to a lack of flushing response to niacin will also, unfortunately, negate any lipid-lowering effect, given that the two share a common receptor.

Because of the negative results of these trials, current guidelines suggest that the use of niacin should be considered only in subjects with triglyceride levels >2.3 mmol/L (200 mg/dL).

31.4.4.3 Triglycerides

The role of elevated triglyceride levels in cardiovascular risk was discussed in more detail elsewhere in the chapter. In summary, recent data suggest that there is a causal link between triglyceride-containing lipoproteins, triglyceride-mediated pathways, and cardiovascular disease. In addition, elevated levels of triglycerides (>10 mmol/l or 880 mg/dl) bring a hugely increased risk of acute pancreatitis and necessitate preventative action. The causes of elevated triglyceride levels are multiple and are summarized in Table 31.7. In cases of secondary hypertriglyceridemia, management should include, where possible, addressing the underlying problem. Several pharmaceutical methods of triglyceride lowering have been identified and their efficacy and effect on cardiovascular risk is discussed next.

TABLE 31.7 Primary and Secondary Causes of Hypertriglyceridemia

Primary	Secondary
Familial hypertriglyceridemia	Alcohol excess
Familial combined hyperlipidemia	Obesity
Familial dysbetalipoproteinemia	Physical inactivity
Familial hypoalphalipoproteinemia	Type 2 diabetes
Familial chylomicronemia	Renal disease
	Pregnancy
	Hypothyroidism
	Autoimmune disorders
	Drugs: thiazide diuretics; β-blockers; estrogens; corticosteroids; isoretinoin; ciclosporin; bile acid-binding resins; antiretroviral protease inhibitors; antipsychotics

Fibrates

Fibrates act to reduce triglyceride levels through the activation of the nuclear transcription factor peroxisome proliferator-activated receptor-α (PPARα). Once activated, PPARα binds to another nuclear receptor, rentinoid X. The resultant heterodimer binds to specific peroxisome proliferator response elements, causing alterations in the expression of genes controlling lipid metabolism. Activation of lipoprotein lipase and apo A-V, along with reduced production of apo C-III, results in increased lipolysis and clearance of triglyceride-rich lipoproteins. Also, the production of Apo-B and VLDL is diminished and the clearance of LDL is increased. Fibrates also have the effect of increasing HDL through increased production of apo A-I/II.

Fibrates have been shown to be effective in lowering triglycerides by up to 50%, with an additional effect on raising HDL-C levels. The latter appears to be greatest in those with low baseline levels. Despite the benefits of fibrate therapy for lipid profiles, the data pointing toward their benefit for cardiovascular outcomes are limited. The Fenofibrate Intervention and Event Lowering in Diabetes (FIELD) study, a large study involving almost 10,000 patients with type 2 diabetes, failed to show any benefit of fibrate versus placebo on the primary end point of cardiovascular event incidence. However, it did show a significant relative reduction (24%) in nonfatal myocardial infarction in its fibrate arm [127].

Similarly, the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial reported no benefit in the combination of fibrate and simvastatin, versus simvastatin alone, on rates of fatal cardiovascular events, nonfatal myocardial infarction, or stroke [128]. A meta-analysis published in 2010 of 18 trials with more than 45,000 patients and 2870 major cardiovascular events, 4552 coronary events, and 3880 deaths, reported a 10% relative risk reduction for major cardiovascular events driven primarily by a 13% relative risk reduction for coronary events, but found no significant benefit on cardiovascular mortality [129]. This benefit was more evident in patients with high baseline triglyceride levels. Current guidelines advocate the use of fibrates to lower triglycerides in patients with severe hypertriglyceremia (>2.3 mmol/L or 200 mg/dL) [24].

Several studies have identified common polymorphisms associated with a variation in triglyceride-lowering in response to fibrate therapy. These include variants at CYP7A1, the APOA1/C3/A4/A5 cluster, and the PPARα locus. In a candidate gene study of loci identified through a large GWAS, in which 95 loci were identified in association with variance in lipid levels, the investigators identified a significant association between increased levels of triglycerides and HDL in response to fibrate therapy and the rs964184 SNP near the APOA1 locus [130].

Omega 3 Fatty Acids

The omega 3 fatty acids (also known as n-3 polyunsaturated fatty acids)—eicosapentenoic acid (EPA), and docosahexenoic acid (DHA)—are naturally occurring fats in fish oils. Administration of fish omega 3 fatty acids as a dietary supplement has been shown to lower triglyceride levels by around 30%, through a mechanism not entirely understood but thought to involve decreased hepatic secretion of VLDL. They have also been purported to inhibit development of atherosclerosis, reduce blood pressure, and have an antiarrhythmic effect. However, there are conflicting results on their beneficial effect on these parameters.

Epidemiological data from the 1970s pointed toward the potentially beneficial effect of omega 3 fatty acids on the CVD risk in Greenland Eskimos. Two large studies have shown a positive benefit, in the role of secondary prevention, in both post-myocardial infarction patients and in patients with heart failure, on CVD mortality and morbidity [131,132]. However, a recently published study demonstrated no benefit associated with omega 3 fatty acids in the primary prevention of death from CVD or in CVD hospital admissions, in a cohort with multiple cardiovascular risk factors, but no history of previous myocardial infarction [133]. Several meta-analyses of trials examining omega 3 fatty acids have reported conflicting results on the effect on CVD risk and mortality [134]. Their use is currently advocated in patients with elevated triglycerides as an adjunct to diet if levels exceed 5.6 mmol/L (496 mg/dl) [24].

31.5 CONCLUSION AND FUTURE DIRECTIONS

This chapter highlighted the important role that lipids play in the body's metabolism and in determining CVD risk. The individual aspects of the plasma lipid profile were examined, highlighting elevated LDL-C levels as the most significant factor in increasing that risk. Measurement of plasma LDL-C levels, along with novel markers such as apo B and Lp(a) remain the mainstay of CVD risk estimation. Observational data have shown low levels of HDL-C to be an independent risk factor; however, pharmacological elevation of HDL-C has, to date, not shown any positive impact on risk reduction. Recent data point toward the role of triglyceride-mediated pathways in increased risk of CVD. This remains an important avenue of investigation.

The identification of monogenic dyslipidemias (e.g., mutations in the LDL receptor, MTP, and PCSK9) has opened up new avenues for pharmacological targeting of lipid pathways. Along with monogenic disorders, multiple common genetic variants have been identified as playing a role in altering plasma lipid levels. However, because the vast majority of these variants have been shown to have relatively small effect sizes, it is thought that the cumulative effect of several variants is more important in determining an individual's plasma lipid levels than individual variants alone. Work is ongoing to identify variants with significant effect sizes in the hope that this will improve cardiovascular risk assessment and potentially highlight new genetic targets for pharmacological therapy.

The extensive work on the pharmacogenetic response to lipid-modifying therapy, in particular statin therapy, has highlighted novel genes and pathways (*SLCO1B1, APOE LDL-R*, *PCSK9*, *KIF6*, *CLMN*) that may enable more targeted management of cardiovascular risk. Current clinical management of dyslipidemia and elevated CVD risk involves administering lipid-lowering therapy to those above a certain threshold of cardiovascular risk estimated from a prognostic model. Monitoring of LDL concentrations is commonly utilized to guide treatment.

There are challenges in implementing pharmacogenomic discoveries in routine clinical practice. Overcoming them requires wider availability of clinically licensed testing, more compelling evidence of efficacy, cost effectiveness analysis, and smoother integration of pharmacogenomic information into clinical decision making. As most statins are now off-patent, there may not be a great incentive to reduce their use. Given that approximately four individuals must be screened to identify one at a higher risk of myopathy, it may well be less expensive do without genotyping and simply start individuals on statin drugs/doses that are less likely to have muscle toxicities.

The example of *SLCO1B1* indicates that pharmacogenomics is a powerful way to stratify patients for risk of adverse effects. It shows that: (1) the genetically mediated risk is not readily identified by other means; (2) the variants are common and the effects are large; and (3) treatment modifications in the form of increased monitoring or alternative therapy are available. Vanderbilt University is an example of an institution where SLCO1B1 genotyping has been implemented into clinical practice [135]. Patients with risk factors for cardiovascular disease who may need statin therapy in the future are genotyped for SLCO1B1 521T>C preemptively and the results are placed in their electronic health record. An electronic order for simvastatin triggers an electronic alert to warn about the increased risk for myopathy in patients with the CC or CT genotype.

REFERENCES

[1] Krauss RM. What can the genome tell us about LDL cholesterol? Lancet 2008;371:450–2.

[2] Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study, 2010. Lancet 2012;380:2224–60.

[3] WHO | Cardiovascular diseases (CVDs). <www.who.int/cardiovascular_diseases/en/> [accessed 13.07.13].

[4] Kannel WB, Castelli WP, Gordon T, McNamara PM. Serum cholesterol, lipoproteins, and the risk of coronary heart disease: the framingham study. Ann Inter Med 1971;74:1–12.

[5] Castelli WP, Anderson K, Wilson PW, Levy D. Lipids and risk of coronary heart disease. The Framingham Study. Ann Epidemiol 1992;2:23–8.

[6] Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, Halsey J, et al. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. Lancet 2007;370:1829–39.

[7] Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet 1994;344:1383–9.

[8] Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG, et al. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and recurrent events trial investigators. N Engl J Med 1996;335:1001–9.

[9] Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, MacFarlane PW, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. N Engl J Med 1995;333:1301–7.

[10] Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet 2005;366:1267–78.

[11] Taylor F, Huffman MD, Macedo AF, Moore THM, Burke M, Davey Smith G, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013;1:CD004816.

[12] Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. Am J Med 1977;62:707–14.

[13] Assmann G, Schulte H, von Eckardstein A, Huang Y. High-density lipoprotein cholesterol as a predictor of coronary heart disease risk. The PROCAM experience and pathophysiological implications for reverse cholesterol transport. Atherosclerosis 1996;124(Suppl):S11–20.

[14] Barter P, Gotto AM, LaRosa JC, Maroni J, Szarek M, Grundy SM, et al. HDL cholesterol, very low levels of LDL cholesterol, and cardiovascular events. N Engl J Med 2007;357:1301–10.

[15] Gordon DJ, Probstfield JL, Garrison RJ, Neaton JD, Castelli WP, Knoke JD, et al. High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies. Circulation 1989;79:8–15.

[16] Sarwar N, Danesh J, Eiriksdottir G, Sigurdsson G, Wareham N, Bingham S, et al. Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. Circulation 2007;115:450–8.

[17] Patel A, Barzi F, Jamrozik K, Lam TH, Ueshima H, Whitlock G, et al. Serum triglycerides as a risk factor for cardiovascular diseases in the Asia-Pacific region. Circulation 2004;110:2678–86.

[18] Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, et al. Major lipids, apolipoproteins, and risk of vascular disease. JAMA 2009;302:1993–2000.

[19] Sarwar N, Sandhu MS, Ricketts SL, Butterworth AS, Di Angelantonio E, Boekholdt SM, et al. Triglyceride-mediated

pathways and coronary disease: collaborative analysis of 101 studies. Lancet 2010;375:1634–9.
[20] Robinson JG, Wang S, Smith BJ, Jacobson TA. Meta-analysis of the relationship between non-high-density lipoprotein cholesterol reduction and coronary heart disease risk. J Am Coll Cardiol 2009;53:316–22.
[21] Sniderman AD, Williams K, Contois JH, Monroe HM, McQueen MJ, de Graaf J, et al. A meta-analysis of low-density lipoprotein cholesterol, non-high-density lipoprotein cholesterol, and apolipoprotein B as markers of cardiovascular risk. Circ Cardiovasc Qual Outcomes 2011;4:337–45.
[22] Erqou S, Kaptoge S, Perry PL, Di Angelantonio E, Thompson A, White IR, et al. Lipoprotein(a) concentration and the risk of coronary heart disease, stroke, and nonvascular mortality. JAMA 2009; 302:412–23.
[23] Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, et al. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. N Engl J Med 2009;361:2518–28.
[24] Reiner Z, Catapano AL, De Backer G, Graham I, Taskinen M-R, Wiklund O, et al. ESC/EAS Guidelines for the management of dyslipidaemias: the Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). Eur Heart J 2011;32:1769–818.
[25] Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. Science 1986;232:34–47.
[26] Rader DJ, Cohen J, Hobbs HH. Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. J Clin Invest 2003;111:1795–803.
[27] Tobert JA. Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. Nat Rev Drug Discov 2003;2:517–26.
[28] Tybjaerg-Hansen A, Humphries SE. Familial defective apolipoprotein B-100: a single mutation that causes hypercholesterolemia and premature coronary artery disease. Atherosclerosis 1992;96:91–107.
[29] Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, Maioli M, et al. Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. Science 2001;292:1394–8.
[30] Arca M, Zuliani G, Wilund K, Campagna F, Fellin R, Bertolini S, et al. Autosomal recessive hypercholesterolaemia in Sardinia, Italy, and mutations in ARH: a clinical and molecular genetic analysis. Lancet 2002;359:841–7.
[31] Abifadel M, Varret M, Rabès J-P, Allard D, Ouguerram K, Devillers M, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet 2003;34:154–6.
[32] Berge KE, Tian H, Graf GA, Yu L, Grishin NV, Schultz J, et al. Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. Science 2000;290:1771–5.
[33] Sharp D, Blinderman L, Combs KA, Kienzle B, Ricci B, Wager-Smith K, et al. Cloning and gene defects in microsomal triglyceride transfer protein associated with abetalipoproteinaemia. Nature 1993;365:65–9.
[34] Tarugi P, Averna M, Di Leo E, Cefalù AB, Noto D, Magnolo L, et al. Molecular diagnosis of hypobetalipoproteinemia: an ENID review. Atherosclerosis 2007;195:e19–27.
[35] Cohen JC, Boerwinkle E, Mosley TH, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med 2006;354:1264–72.
[36] Sviridov D, Nestel PJ. Genetic factors affecting HDL levels, structure, metabolism and function. Curr Opin Lipidol 2007;18:157–63.
[37] Sirtori CR, Calabresi L, Franceschini G, Baldassarre D, Amato M, Johansson J, et al. Cardiovascular status of carriers of the apolipoprotein A-I(Milano) mutant: the Limone sul Garda study. Circulation 2001;103:1949–54.
[38] Rust S, Rosier M, Funke H, Real J, Amoura Z, Piette JC, et al. Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. Nat Genet 1999;22:352–5.
[39] Hobbs HH, Rader DJ. ABC1: connecting yellow tonsils, neuropathy, and very low HDL. J Clin Invest 1999;104:1015–7.
[40] Kuivenhoven JA, Pritchard H, Hill J, Frohlich J, Assmann G, Kastelein J. The molecular pathology of lecithin:cholesterol acyltransferase (LCAT) deficiency syndromes. J Lipid Res 1997;38:191–205.
[41] Inazu A, Brown ML, Hesler CB, Agellon LB, Koizumi J, Takata K, et al. Increased high-density lipoprotein levels caused by a common cholesteryl-ester transfer protein gene mutation. N Engl J Med 1990; 323:1234–8.
[42] Barter PJ, Brewer HB, Chapman MJ, Hennekens CH, Rader DJ, Tall AR. Cholesteryl ester transfer protein: a novel target for raising HDL and inhibiting atherosclerosis. Arterioscler Thromb Vasc Biol 2003;23:160–7.
[43] Gehrisch S. Common mutations of the lipoprotein lipase gene and their clinical significance. Current Atheroscler Rep 1999;1:70–8.
[44] Walden CC, Hegele RA. Apolipoprotein E in hyperlipidemia. Ann Intern Med 1994;120:1026–36.
[45] Mahley RW, Huang Y, Rall SC. Pathogenesis of type III hyperlipoproteinemia (dysbetalipoproteinemia). Questions, quandaries, and paradoxes. J Lipid Res 1999;40:1933–49.
[46] Hegele RA, Vezina C, Moorjani S, Lupien PJ, Gagne C, Brun LD, et al. A hepatic lipase gene mutation associated with heritable lipolytic deficiency. J Clin Endocrinol Metab 1991;72:730–2.
[47] Ruel IL, Couture P, Gagne C, Deshaies Y, Simard J, Hegele RA, et al. Characterization of a novel mutation causing hepatic lipase deficiency among French Canadians. J Lipid Res 2003;44:1508–14.
[48] Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707–13.
[49] Asselbergs FW, Guo Y, van Iperen EPA, Sivapalaratnam S, Tragante V, Lanktree MB, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. Am J Hum Genet 2012;91:823–38.
[50] Shah S, Casas JP, Gaunt TR, Cooper J, Drenos F, Zabaneh D, et al. Influence of common genetic variation on blood lipid levels, cardiovascular risk, and coronary events in two British prospective cohort studies. Eur Heart J 2013;34:972–81.
[51] Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 2010;466:714–9.
[52] Linsel-Nitschke P, Heeren J, Aherrahrou Z, Bruse P, Gieger C, Illig T, et al. Genetic variation at chromosome 1p13.3 affects sortilin mRNA expression, cellular LDL-uptake and serum LDL levels which translates to the risk of coronary artery disease. Atherosclerosis 2010;208:183–9.
[53] Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 2009;41:334–41.
[54] Kjolby M, Andersen OM, Breiderhoff T, Fjorback AW, Pedersen KM, Madsen P, et al. Sort1, encoded by the cardiovascular risk

locus 1p13.3, is a regulator of hepatic lipoprotein export. Cell Metab 2010;12:213–23.

[55] Qasim A, Rader DJ. Human genetics of variation in high-density lipoprotein cholesterol. Current Atheroscler Rep 2006;8:198–205.

[56] Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet 2012;380:572–80.

[57] Van Capelleveen JC, Bochem AE, Motazacker MM, Hovingh GK, Kastelein JJP. Genetics of HDL-C: a causal link to atherosclerosis? Current Atheroscler Rep 2013;15:326.

[58] Cefalù AB, Noto D, Arpi ML, Yin F, Spina R, Hilden H, et al. Novel LMF1 nonsense mutation in a patient with severe hypertriglyceridemia. J Clin Endocrinol Metab 2009;94:4584–90.

[59] Wang J, Hegele RA. Homozygous missense mutation (G56R) in glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 (GPI-HBP1) in two siblings with fasting chylomicronemia (MIM 144650). Lipids Health Dis 2007;6:23.

[60] Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat Genet 2010;42:684–7.

[61] Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhala N, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet 2010;376:1670–81.

[62] Naci H, Brugts JJ, Fleurence R, Ades A. Dose-comparative effects of different statins on serum lipid levels: a network meta-analysis of 256,827 individuals in 181 randomized controlled trials. Eur J Prev Cardiol 2013;20:658–70.

[63] Knopp RH. Drug treatment lipid disorder. N Engl J Med 1999;341:498–511.

[64] Norman DJ, Illingworth DR, Munson J, Hosenpud J. Myolysis and acute renal failure in a heart-transplant recipient receiving lovastatin. N Engl J Med 1988;318:46–7.

[65] How a statin might destroy a drug company. Lancet 2003;361:793.

[66] Gruer PJ, Vega JM, Mercuri MF, Dobrinska MR, Tobert JA. Concomitant use of cytochrome P450 3A4 inhibitors and simvastatin. Am J Cardiol 1999;84:811–5.

[67] Smith PF, Eydelloth RS, Grossman SJ, Stubbs RJ, Schwartz MS, Germershausen JI, et al. HMG-CoA reductase inhibitor-induced myopathy in the rat: cyclosporine A interaction and mechanism studies. J Pharmacol Exp Ther 1991;257:1225–35.

[68] Brunham LR, Lansberg PJ, Zhang L, Miao F, Carter C, Hovingh GK, et al. Differential effect of the rs4149056 variant in SLCO1B1 on myopathy associated with simvastatin and atorvastatin. Pharmacogenomics J 2012;12:233–7.

[69] Niemi M, Pasanen MK, Neuvonen PJ. Organic anion transporting polypeptide 1B1: a genetically polymorphic transporter of major importance for hepatic drug uptake. Pharmacol Rev 2011;63:157–81.

[70] Maggo SDS, Kennedy MA, Clark DWJ. Clinical implications of pharmacogenetic variation on the effects of statins. Drug Saf Int J Med Toxicol Drug Exp 2011;34:1–19.

[71] Wilke RA, Ramsey LB, Johnson SG, Maxwell WD, McLeod HL, Voora D, et al. The clinical pharmacogenomics implementation consortium: CPIC guideline for SLCO1B1 and simvastatin-induced myopathy. Clin Pharmacol Ther 2012;92:112–7.

[72] Ridker PM, Danielson E, Fonseca FAH, Genest J, Gotto AM, Kastelein JJP, et al. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. N Engl J Med 2008;359:2195–207.

[73] Freeman DJ, Norrie J, Sattar N, Neely RD, Cobbe SM, Ford I, et al. Pravastatin and the development of diabetes mellitus: evidence for a protective treatment effect in the West of Scotland Coronary Prevention Study. Circulation 2001;103:357–62.

[74] Sattar N, Preiss D, Murray HM, Welsh P, Buckley BM, de Craen AJM, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. Lancet 2010;375:735–42.

[75] Carter AA, Gomes T, Camacho X, Juurlink DN, Shah BR, Mamdani MM. Risk of incident diabetes among patients treated with statins: population based study. BMJ (Clinical research ed) 2013;346:f2610.

[76] Bakhai A, Rigney U, Hollis S, Emmas C. Co-administration of statins with cytochrome P450 3A4 inhibitors in a UK primary care population. Pharmacoepidemiol Drug Saf 2012;21:485–93.

[77] Hopewell JC, Parish S, Offer A, Link E, Clarke R, Lathrop M, et al. Impact of common genetic variation on response to simvastatin therapy among 18 705 participants in the Heart Protection Study. Eur Heart J 2013;34:982–92.

[78] Kajinami K, Brousseau ME, Ordovas JM, Schaefer EJ. CYP3A4 genotypes and plasma lipoprotein levels before and after treatment with atorvastatin in primary hypercholesterolemia. Am J Cardiol 2004;93:104–7.

[79] Kivistö KT, Niemi M, Schaeffeler E, Pitkälä K, Tilvis R, Fromm MF, et al. Lipid-lowering response to statins is affected by CYP3A5 polymorphism. Pharmacogenetics 2004;14:523–5.

[80] Mulder AB, van Lijf HJ, Bon MA, van den Bergh FA, Touw DJ, Neef C, et al. Association of polymorphism in the cytochrome CYP2D6 and the efficacy and tolerability of simvastatin. Clin Pharmacol Ther 2001;70:546–51.

[81] Frudakis TN, Thomas MJ, Ginjupalli SN, Handelin B, Gabriel R, Gomez HJ. CYP2D6*4 polymorphism is associated with statin-induced muscle effects. Pharmacogenet Genomics 2007;17:695–707.

[82] Takane H, Miyata M, Burioka N, Shigemasa C, Shimizu E, Otsubo K, et al. Pharmacogenetic determinants of variability in lipid-lowering response to pravastatin therapy. J Hum Genet 2006;51:822–6.

[83] Chasman DI, Posada D, Subrahmanyan L, Cook NR, Stanton VP, Ridker PM. Pharmacogenetic study of statin therapy and cholesterol reduction. JAMA 2004;291:2821–7.

[84] Donnelly LA, Doney ASF, Dannfald J, Whitley AL, Lang CC, Morris AD, et al. A paucimorphic variant in the HMG-CoA reductase gene is associated with lipid-lowering response to statin treatment in diabetes: a GoDARTS study. Pharmacogenet Genomics 2008;18:1021–6.

[85] Kajinami K, Brousseau ME, Nartsupha C, Ordovas JM, Schaefer EJ. ATP binding cassette transporter G5 and G8 genotypes and plasma lipoprotein levels before and after treatment with atorvastatin. J Lipid Res 2004;45:653–6.

[86] Donnelly LA, Palmer CNA, Whitley AL, Lang CC, Doney ASF, Morris AD, et al. Apolipoprotein E genotypes are associated with lipid-lowering responses to statin treatment in diabetes: a Go-DARTS study. Pharmacogenet Genomics 2008;18:279–87.

[87] Gerdes LU, Gerdes C, Kervinen K, Savolainen M, Klausen IC, Hansen PS, et al. The apolipoprotein epsilon4 allele determines prognosis and the effect on prognosis of simvastatin in survivors of myocardial infarction : a substudy of the Scandinavian simvastatin survival study. Circulation 2000;101:1366–71.

[88] Maitland-van der Zee A-H, Jukema JW, Zwinderman AH, Hallman DM, De Boer A, Kastelein JJP, et al. Apolipoprotein-E polymorphism and response to pravastatin in men with coronary artery disease (REGRESS). Acta Cardiol 2006;61:327–31.

[89] Regieli JJ, Jukema JW, Grobbee DE, Kastelein JJP, Kuivenhoven JA, Zwinderman AH, et al. CETP genotype predicts increased mortality in statin-treated men with proven cardiovascular disease: an adverse pharmacogenetic interaction. Eur Heart J 2008;29:2792–9.

[90] Iakoubova OA, Tong CH, Rowland CM, Kirchgessner TG, Young BA, Arellano AR, et al. Association of the Trp719Arg polymorphism in kinesin-like protein 6 with myocardial infarction and coronary heart disease in 2 prospective trials: the CARE and WOSCOPS trials. J Am Coll Cardiol 2008;51:435–43.

[91] Assimes TL, Hólm H, Kathiresan S, Reilly MP, Thorleifsson G, Voight BF, et al. Lack of association between the Trp719Arg polymorphism in kinesin-like protein-6 and coronary artery disease in 19 case-control studies. J Am Coll Cardiol 2010;56:1552–63.

[92] Hopewell JC, Parish S, Clarke R, Armitage J, Bowman L, Hager J, et al. No impact of KIF6 genotype on vascular risk and statin response among 18,348 randomized patients in the heart protection study. J Am Coll Cardiol 2011;57:2000–7.

[93] Pandor A, Ara RM, Tumur I, Wilkinson AJ, Paisley S, Duenas A, et al. Ezetimibe monotherapy for cholesterol lowering in 2,722 people: systematic review and meta-analysis of randomized controlled trials. J Intern Med 2009;265:568–80.

[94] Ballantyne CM, Houri J, Notarbartolo A, Melani L, Lipka LJ, Suresh R, et al. Effect of ezetimibe coadministered with atorvastatin in 628 patients with primary hypercholesterolemia: a prospective, randomized, double-blind trial. Circulation 2003;107:2409–15.

[95] Davidson MH, McGarry T, Bettis R, Melani L, Lipka LJ, LeBeaut AP, et al. Ezetimibe coadministered with simvastatin in patients with primary hypercholesterolemia. J Am Coll Cardiol 2002;40:2125–34.

[96] Kastelein JJP, Akdim F, Stroes ESG, Zwinderman AH, Bots ML, Stalenhoef AFH, et al. Simvastatin with or without ezetimibe in familial hypercholesterolemia. N Engl J Med 2008;358:1431–43.

[97] Baigent C, Landray MJ, Reith C, Emberson J, Wheeler DC, Tomson C, et al. The effects of lowering LDL cholesterol with simvastatin plus ezetimibe in patients with chronic kidney disease (Study of heart and renal protection): a randomised placebo-controlled trial. Lancet 2011;377:2181–92.

[98] Peto R, Emberson J, Landray M, Baigent C, Collins R, Clare R, et al. Analyses of cancer data from three ezetimibe trials. N Engl J Med 2008;359:1357–66.

[99] Wang J, Williams CM, Hegele RA. Compound heterozygosity for two non-synonymous polymorphisms in NPC1L1 in a non-responder to ezetimibe. Clin Genet 2005;67:175–7.

[100] Hegele RA, Guy J, Ban MR, Wang J. NPC1L1 haplotype is associated with inter-individual variation in plasma low-density lipoprotein response to ezetimibe. Lipids Health Dis 2005;4:16.

[101] Fonseca VA, Handelsman Y, Staels B. Colesevelam lowers glucose and lipid levels in type 2 diabetes: the clinical evidence. Diabetes Obes Metab 2010;12:384–92.

[102] The lipid research clinics coronary primary prevention trial results: I. Reduction in incidence of coronary heart disease. JAMA 1984;251:351.

[103] Huijgen R, Abbink EJ, Bruckert E, Stalenhoef AFH, Imholz BPM, Durrington PN, et al. Colesevelam added to combination therapy with a statin and ezetimibe in patients with familial hypercholesterolemia: a 12-week, multicenter, randomized, double-blind, controlled trial. Clin Ther 2010;32:615–25.

[104] Santos RD, Waters DD, Tarasenko L, Messig M, Jukema JW, Chiang C-W, et al. A comparison of non-HDL and LDL cholesterol goal attainment in a large, multinational patient population: the Lipid Treatment Assessment Project 2. Atherosclerosis 2012;224:150–3.

[105] Cuchel M, Bloedon LT, Szapary PO, Kolansky DM, Wolfe ML, Sarkis A, et al. Inhibition of microsomal triglyceride transfer protein in familial hypercholesterolemia. N Engl J Med 2007;356:148–56.

[106] Cuchel M, Meagher EA, du Toit Theron H, Blom DJ, Marais AD, Hegele RA, et al. Efficacy and safety of a microsomal triglyceride transfer protein inhibitor in patients with homozygous familial hypercholesterolaemia: a single-arm, open-label, phase 3 study. Lancet 2013;381:40–6.

[107] Stein EA, Dufour R, Gagne C, Gaudet D, East C, Donovan JM, et al. Apolipoprotein B synthesis inhibition with mipomersen in heterozygous familial hypercholesterolemia: results of a randomized, double-blind, placebo-controlled trial to assess efficacy and safety as add-on therapy in patients with coronary artery disease. Circulation 2012;126:2283–92.

[108] Visser ME, Wagener G, Baker BF, Geary RS, Donovan JM, Beuers UHW, et al. Mipomersen, an apolipoprotein B synthesis inhibitor, lowers low-density lipoprotein cholesterol in high-risk statin-intolerant patients: a randomized, double-blind, placebo-controlled trial. Eur Heart J 2012;33:1142–9.

[109] Maxwell KN, Breslow JL. Adenoviral-mediated expression of Pcsk9 in mice results in a low-density lipoprotein receptor knockout phenotype. Proc Nat Acad Sci USA 2004;101:7100–5.

[110] Graham MJ, Lemonidis KM, Whipple CP, Subramaniam A, Monia BP, Crooke ST, et al. Antisense inhibition of proprotein convertase subtilisin/kexin type 9 reduces serum LDL in hyperlipidemic mice. J Lipid Res 2007;48:763–7.

[111] Stein EA, Gipe D, Bergeron J, Gaudet D, Weiss R, Dufour R, et al. Effect of a monoclonal antibody to PCSK9, REGN727/SAR236553, to reduce low-density lipoprotein cholesterol in patients with heterozygous familial hypercholesterolaemia on stable statin dose with or without ezetimibe therapy: a phase 2 randomised controlle. Lancet 2012;380:29–36.

[112] Roth EM, McKenney JM, Hanotin C, Asset G, Stein EA. Atorvastatin with or without an antibody to PCSK9 in primary hypercholesterolemia. N Engl J Med 2012;367:1891–900.

[113] Ridker PM, Paré G, Parker AN, Zee RYL, Miletich JP, Chasman DI. Polymorphism in the CETP gene region, HDL cholesterol, and risk of future myocardial infarction: genomewide analysis among 18 245 initially healthy women from the Women's Genome Health Study. Circ Cardiovas Genetics 2009;2:26–33.

[114] Thompson A, Di Angelantonio E, Sarwar N, Erqou S, Saleheen D, Dullaart RPF, et al. Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. JAMA 2008;299:2777–88.

[115] Brousseau ME, Schaefer EJ, Wolfe ML, Bloedon LT, Digenio AG, Clark RW, et al. Effects of an inhibitor of cholesteryl ester transfer protein on HDL cholesterol. N Engl J Med 2004;350:1505–15.

[116] Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJP, Komajda M, et al. Effects of torcetrapib in patients at high risk for coronary events. N Engl J Med 2007;357:2109–22.

[117] Lüscher TF, Taddei S, Kaski J-C, Jukema JW, Kallend D, Münzel T, et al. Vascular effects and safety of dalcetrapib in patients with or at risk of coronary heart disease: the dal-VESSEL randomized clinical trial. Eur Heart J 2012;33:857–65.

[118] Fayad ZA, Mani V, Woodward M, Kallend D, Abt M, Burgess T, et al. Safety and efficacy of dalcetrapib on atherosclerotic disease

using novel non-invasive multimodality imaging (dal-PLAQUE): a randomised clinical trial. Lancet 2011;378:1547–59.

[119] Schwartz GG, Olsson AG, Abt M, Ballantyne CM, Barter PJ, Brumm J, et al. Effects of dalcetrapib in patients with a recent acute coronary syndrome. N Engl J Med 2012;367:2089–99.

[120] Cannon CP, Shah S, Dansky HM, Davidson M, Brinton EA, Gotto AM, et al. Safety of anacetrapib in patients with or at high risk for coronary heart disease. N Engl J Med 2010;363:2406–15.

[121] Nicholls SJ, Brewer HB, Kastelein JJP, Krueger KA, Wang M-D, Shao M, et al. Effects of the CETP inhibitor evacetrapib administered as monotherapy or in combination with statins on HDL and LDL cholesterol: a randomized controlled trial. JAMA 2011;306:2099–109.

[122] Zellner C, Pullinger CR, Aouizerat BE, Frost PH, Kwok P-Y, Malloy MJ, et al. Variations in human HM74 (GPR109B) and HM74A (GPR109A) niacin receptors. Hum Mutat 2005;25:18–21.

[123] Canner PL, Berge KG, Wenger NK, Stamler J, Friedman L, Prineas RJ, et al. Fifteen year mortality in Coronary Drug Project patients: long-term benefit with niacin. J Am Coll Cardiol 1986;8:1245–55.

[124] Taylor AJ, Lee HJ, Sullenberger LE. The effect of 24 months of combination statin and extended-release niacin on carotid intima-media thickness: ARBITER 3. Curr Med Res Opin 2006;22:2243–50.

[125] Boden WE, Probstfield JL, Anderson T, Chaitman BR, Desvignes-Nickens P, Koprowicz K, et al. Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. N Engl J Med 2011;365:2255–67.

[126] Benyó Z, Gille A, Kero J, Csiky M, Suchánková MC, Nüsing RM, et al. GPR109A (PUMA-G/HM74A) mediates nicotinic acid-induced flushing. J Clin Invest 2005;115:3634–40.

[127] Keech A, Simes RJ, Barter P, Best J, Scott R, Taskinen MR, et al. Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. Lancet 2005;366:1849–61.

[128] Ginsberg HN, Elam MB, Lovato LC, Crouse JR, Leiter LA, Linz P, et al. Effects of combination lipid therapy in type 2 diabetes mellitus. N Engl J Med 2010;362:1563–74.

[129] Jun M, Foote C, Lv J, Neal B, Patel A, Nicholls SJ, et al. Effects of fibrates on cardiovascular outcomes: a systematic review and meta-analysis. Lancet 2010;375:1875–84.

[130] Aslibekyan S, Goodarzi MO, Frazier-Wood AC, Yan X, Irvin MR, Kim E, et al. Variants identified in a GWAS meta-analysis for blood lipids are associated with the lipid response to fenofibrate. PLoS ONE 2012;7:e48663.

[131] Tavazzi L, Maggioni AP, Marchioli R, Barlera S, Franzosi MG, Latini R, et al. Effect of n-3 polyunsaturated fatty acids in patients with chronic heart failure (the GISSI-HF trial): a randomised, double-blind, placebo-controlled trial. Lancet 2008;372:1223–30.

[132] Dietary supplementation with n-3 polyunsaturated fatty acids and vitamin E after myocardial infarction: results of the GISSI-Prevenzione trial. Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto miocardico. Lancet 1999;354:447–55.

[133] Roncaglioni MC, Tombesi M, Avanzini F, Barlera S, Caimi V, Longoni P, et al. N-3 fatty acids in patients with multiple cardiovascular risk factors. N Engl J Med 2013;368:1800–8.

[134] Lorente-Cebrián S, Costa AGV, Navas-Carretero S, Zabala M, Martínez JA, Moreno-Aliaga MJ. Role of omega-3 fatty acids in obesity, metabolic syndrome, and cardiovascular diseases: a review of the evidence. J Physiol Biochem 2013;69(3):633–51. http://dx.doi.org/10.1007/s13105-013-0265-4.

[135] Lieb W, Völzke H, Pulley JM, Roden DM, Kroemer HK. Strategies for personalized medicine-based research and implementation in the clinical workflow. Clin Pharmacol Ther 2012;92:443–5.

Chapter 32

Hypertension Pharmacogenomics

Yan Gong,[a] Caitrin W. McDonough,[a] Sandosh Padmanabhan,[b] and Julie A Johnson[a,c]
[a]*Department of Pharmacotherapy and Translational Research, Center for Pharmacogenomics, University of Florida, USA,* [b]*BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK,* [c]*Department of Medicine, University of Florida, USA*

Chapter Outline

32.1 Overview **747**
32.2 Epidemiology and Public Health Impact **747**
  32.2.1 Risk Factors for Hypertension 748
  32.2.2 Hypertension Causation 748
  32.2.3 Heritability of Blood Pressure and Hypertension 748
32.3 Genetic Basis of Hypertension **749**
  32.3.1 Sodium and Vascular Volume 749
  32.3.2 Autonomic Nervous System 753
  32.3.3 Vascular Mechanisms 753
32.4 Need for Stratified Approaches to Treatment with Antihypertensives **754**
  32.4.1 Nongenetic Approaches 755
  32.4.2 Genetic Approaches 758
32.5 Pharmacogenomics of Blood Pressure Response to Antihypertensive Drugs **758**
  32.5.1 Thiazide Diuretics 758
  32.5.2 β-Blockers 761
  32.5.3 Calcium Channel Blockers 762
  32.5.4 ACE Inhibitors and ARBs 763
32.6 Pharmacogenomics of Antihypertensive Treatment—Related Outcomes **764**
  32.6.1 Clinical Trials 764
  32.6.2 Genes with Strong Evidence of Association with Treatment-Related Outcomes 764
  32.6.3 Additional Genes with Evidence of Association with Treatment-Related Outcomes 769
32.7 Future Perspectives **771**
References **772**

32.1 OVERVIEW

Blood pressure (BP) is determined by cardiac output and peripheral vascular resistance—cardiac output depends on stroke volume and heart rate, while peripheral resistance depends on functional and anatomic changes in the small arteries (lumen diameter 100–400 μm) and arterioles. BP is expressed as two numbers—systolic BP (SBP) is the highest pressure in each cardiac cycle and is related to cardiac output, while diastolic BP (DBP) is the lowest pressure in the cardiac cycle and reflects peripheral vascular resistance. BP is a quantitative trait and is normally distributed in the general population. Variation in extracellular fluid volume, the contractile state of the heart, and vascular tone determine variation in BP level.

32.2 EPIDEMIOLOGY AND PUBLIC HEALTH IMPACT

In adults, there is a continuous, incremental risk of cardiovascular disease, stroke, and renal disease, with increasing BP starting as low as 115/75 mmHg [1]. Hypertension as a disease process was borne out in epidemiologic studies demonstrating that people with elevated BP have a greater risk of developing strokes and heart attacks than matched people with normal BP. Hypertension is defined based on a cut-off at the upper end of BP distribution, “at which the benefits of action (i.e., therapeutic intervention) exceed those of inaction” [2]. Based on this definition, there are more than 1 billion people with high BP worldwide, and the World Health Organization (WHO) suggests that this will

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00032-3
© 2014 Elsevier Inc. All rights reserved.

rise to 1.5 billion by 2020 [3]. Hypertension is estimated to contribute 3.5 times more to the total global disease burden of cardiovascular disease than smoking and 1.6 times more than hypercholesterolemia. Worldwide, 20% of deaths in men, 24% of deaths in women, 62% of strokes, and 49% of heart disease are attributable to BP [4].

32.2.1 Risk Factors for Hypertension

The risk factors for hypertension can be classified into modifiable (weight, diet, smoking, alcohol intake, diabetes, physical activity, psychosocial stressors) and unmodifiable factors (age, sex, ethnicity, family history). Males generally have higher BP than females. SBP increases linearly at age 30–84 years along with mean arterial pressure, but DBP increases linearly up to age 50–60 years, after which it begins to decline, with a steep increase in pulse pressure [5,6]. The late decline of DBP after the age of 60 years and the continuous rise in SBP reflect increased large-artery stiffness in older age. Positive associations have also been documented between body mass index (BMI) and BP in both cross-sectional and prospective studies [7]. The odds of progression to hypertension increase by 20–30% for every 5% gain in body weight. There is a consistent dose–response relation between salt intake and BP within the range of 3–12 g salt per day. A reduction of only 3 g per day predicts a decrease in BP of 3.6–5.6/1.9–3.2 mmHg in hypertensive subjects and 1.8–3.5/0.8–1.8 mmHg in normotensive subjects [8,9]. At all ages, hypertension is more common in African-Americans than in whites; in all ethnic and racial groups, it is more common in the economically disadvantaged. Interestingly, hypertension is more prevalent in the African-American population in the United States than in African Caribbean and native black African populations [10,11]. In some societies, BP shows only a small age-related increase and may be related in part to an agrarian lifestyle and in part to the high-potassium and low-sodium diet of the hunter-gatherer, the low level of formal education, a more rural lifestyle, a lower consumption of food, and cultural isolation [12–15].

From an evolutionary perspective, high BP may be the undesirable pleiotropic effect of a genotype that may have optimized fitness in an ancient environment. When humans arrived >40,000 years ago, they lived in this primarily salt-poor environment consuming <690 mg of salt per day. This required ancestral humans to develop a kidney-based regulatory mechanism for the conservation of extracellular fluid and electrolytes; natural selection over generations molded their adaptation to the low salt content in their environment. However, in modern society access to salt is almost unlimited, with a sodium consumption of about 4 g or more per day (equivalent to 10 g of salt) [16]. Thus, high BP is a disease of modern civilization with its abundance of dietary salt, processed foods, and longer lifespans [17]. Different populations show different predispositions to the trait, and this may reflect different evolutionary selection pressures given that different populations do not necessarily share the same ancestral histories. For example, hypertension occurs earlier and with more severity in people of African ancestry compared to those of European ancestry [18]. The rates of hypertension and sodium sensitivity are generally higher in individuals carrying the ancestral alleles of sodium-conserving genes, which show strong latitudinal clines with the ancestral sodium-conserving alleles more prevalent in African populations and less so in northern regions [19–22]. It is also hypothesized that the renin–angiotensin–aldosterone system (RAAS) was initially adapted for sodium conservation and may be maladaptive in modern societies with ready access to dietary sodium [22].

32.2.2 Hypertension Causation

The majority (95%) of individuals with hypertension have essential or primary hypertension, in which no single underlying cause is identified for the high BP, which is thought to result from a genetic predisposition underlying the cumulative effects of various lifestyle factors over many years. The remaining 5% of patients have secondary hypertension, which is due to an underlying disease or is a side effect of medication. The most common underlying condition is chronic kidney disease, but others include monogenic syndromes and structural abnormalities in the cardiovascular system.

32.2.3 Heritability of Blood Pressure and Hypertension

Family studies have consistently demonstrated a genetic component influencing BP and hypertension. The Montreal adoption study [23] demonstrated correlation coefficients of 0.38 and 0.16 between biological and adoptive siblings, respectively, while the Victorian Family Heart Study [24] estimated correlation coefficients of 0.44 for nontwin siblings, 0.78 for monozygous twins, 0.50 for dizygous twins, and 0.12 for spouse–spouse pairs. Hunt et al. [25] showed familial aggregation of hypertension from an analysis of life table data for 94,292 persons, and found the relative risks of developing hypertension to be 4.1 in men and 5 in women aged 20–39 who had at least two first-degree relatives affected by hypertension. Two additional measures that are commonly used to assess the genetic component of a trait are heritability (h^2)—the fraction of variation in disease susceptibility due to genetic factors, and sibling recurrent risk (λs)—the degree of elevated risk of disease for a sibling of an affected individual compared with a member of the general population. The heritability of clinical SBP and DBP is around 15–40% and 15–30%, respectively; for ambulatory night-time SBP and DBP, the heritabilities are 32–70% and 32–50% [26–32]. The sibling recurrent risk

(λ_s) of hypertension is ~1.2–1.5 [33]. Taking this along with heritability and correlation estimates, hypertension and BP can be considered traits with relatively modest genetic effects.

32.3 GENETIC BASIS OF HYPERTENSION

The genetic basis of hypertension is best understood by reviewing the causative genes for the rare monogenic forms of hypertension and the common single-nucleotide polymorphisms (SNPs) that have been identified from large-scale genome-wide association studies (GWASs) (Table 32.1, Figure 32.1).

32.3.1 Sodium and Vascular Volume

Vascular volume is a primary determinant of arterial pressure over the long term; sodium is the principal determinant of extracellular fluid volume. The effect of sodium on BP is related to NaCl—nonchloride sodium salts have little or no effect on BP [34,35]. Mendelian forms of syndromic hypotension and hypertension [36] have all been linked to mutations in genes whose encoded proteins regulate the balance of salt and water in the kidney (Table 32.1).

Glucocorticoid-remediable aldosteronism or familial hyperaldosteronism type 1 (FH-I, OMIM #103900) is an autosomal dominant syndrome in which hypertension is caused by increased aldosterone secretion driven by

TABLE 32.1 Genetic Loci Associated with Monogenic BP Syndromes and Identified Through GWAS

CHR	Gene/Nearest Gene	Syndrome	GWAS	Pathway	Antihypertensive Therapy
1p36.13	*CLCNKB*	Bartter syndrome, type 3 OMIM #607364		Renal electrolyte balance	Low/normal BP
11q24.3	*KCNJ1*	Bartter syndrome, antenatal, type 2 OMIM #241200		Renal electrolyte balance	Aldosterone antagonists, spironolactone, ACEI, Indomethacin
15q21.1	*SLC12A1*	Bartter syndrome, antenatal, type 1 OMIM #601678		Renal electrolyte balance	Potassium chloride supplementation Potassium-sparing diuretics (e.g., spironolactone, amilioride, triamterene) Indomethacin ACE inhibitors
2q36.2	*CUL3*	Pseudohypoaldosteronism, type IIE OMIM *603136		Renal electrolyte balance	Dietary Na restriction, hydrochlorothiazide
17q21.3	*WNK4*	Pseudohypoaldosteronism type IIB Gordon's syndrome		Renal electrolyte balance	Dietary Na restriction, hydrochlorothiazide
4q31.2	*NR3C2*	Hypertension exacerbation in pregnancy OMIM #605115		Autosomal dominant Missense mutation (S810L) in mineralocorticoid receptor low-renin, low-aldosteronehypokalemia	
		Pseudohypoaldosteronism type I OMIM #177735		Renal electrolyte balance	High-Na, low-K diet, K-binding resins, hydrochlorothiazide
5q31.2	*KLHL3*	Pseudohypoaldosteronism, type IID OMIM #614495		Renal electrolyte balance	Dietary Na restriction, hydrochlorothiazide
12p12.3	*WNK1*	Pseudohypoaldosteronism type IIC Gordon's syndrome		Renal electrolyte balance	Dietary Na restriction, hydrochlorothiazide

(*Continued*)

TABLE 32.1 Continued

CHR	Gene/Nearest Gene	Syndrome	GWAS	Pathway	Antihypertensive Therapy
16p12.2	*SCNN1B, SCNN1G*	Liddle syndrome OMIM #177200		Renal electrolyte balance	Amiloride or triamterene
16q13	*SLC12A3*	Gitelman syndrome OMIM #263800		Renal electrolyte balance	Potassium and magnesium supplements amiloride, spironolactone, triamterene
7p22		Familial hyperaldosteronism type 2 OMIM #605635		Steroid/aldosterone synthesis	Spironolactone Not suppressed by dexamethasone
8q24.3	*CYP11B1, CYP11B2*	Familial hyperaldosteronism type 1 Glucocorticoid remediable aldosteronism (GRA) OMIM #103900			Low-dose glucocorticoids, amiloride, spironolactone
8q24.3	*CYP11B2*	Corticosterone methyloxidase II deficiency OMIM #61060			Mineralocorticoid replacement
8q24.3	*CYP11B1*	Steroid 11β-hydroxylase deficiency OMIM #202010			Glucocorticoid replacement, standard antihypertensive drugs
10q24.3	*CYP17A1*	17 α-hydroxylase and/or 17,20-lyase deficiency	CHARGE, GBPG, AGEN-BP, ICBP		Glucocorticoid therapy
16q22.1	*HSD11B2*	Apparent mineralocorticoid excess OMIM # 218030			Spironolactone Dexamethasone
12p12.2		Hypertension with brachydactyly Bilginturan syndrome			Standard antihypertensive drugs
1p36.13	*SDHB*	Paragangliomas 4 OMIM #115310		Sympathetic system	Treatment for pheochromocytoma-alpha-beta-adrenergic blockade followed by surgery
1q23.3	*SDHC*	Paragangliomas 3 OMIM #605373			
3p25.3	*VHL*	von Hippel-Lindau syndrome OMIM #193300			
5p15.3	*SDHA*	Paragangliomas 5 OMIM # 614165			
10q11.2	*RET*	Multiple endocrine neoplasia type IIA OMIM #171400			
11q12.2	*SDHAF2*	Paragangliomas 2 OMIM #601650			
11q23.1	*SDHD*	Paragangliomas 1 OMIM #16800			

TABLE 32.1 Continued

CHR	Gene/Nearest Gene	Syndrome	GWAS	Pathway	Antihypertensive Therapy
1p36.2	*MTHFR (NPPA, NPPB)*		CHARGE, GBPG, AGEN, ICBP Gene-centric	Neprilysin or neutral endopeptidase that increases A- and B-type natriuretic peptide concentrations (in development)	Neprilysin or neutral endopeptidase that increase A- and B-type natriuretic peptide concentrations (in development)
1q42.2	*AGT*		Gene-centric	Renal electrolyte balance, vascular function	
3q22.1	*ULK4*		CHARGE, GBPG, ICBP		
3q26.2	*MECOM (MDS1)*		GBPG, ICBP		
4q21.2	*FGF5*		GBPG, AGEN, ICBP		
5p13.3	*NPR3*		AGEN, ICBP Gene-centric	Natriuretic peptide clearance receptor	
6p22.2	*HFE*	Hemochromatosis OMIM #235200	ICBP Gene-centric		
7q36.1	*NOS3*	Pregnancy-induced hypertension OMIM +163729	HYPERGENES, Gene-centric	Endothelial function	
10p12.3	*CACNB2*		CHARGE, ICBP	Vascular/cardiac function?	
11p15.1	*PLEKHA7*		CHARGE, ICBP		
11p15.2	*SOX6*		Gene-centric	Transcription	
11p15.5	*LSP1/TNNT3*		Gene-centric	Endothelial function?	
12q21.3	*ATP2B1*		CHARGE, GBPG, AGEN, ICBP Gene-centric	Vascular function?	
12q24.1	*SH2B3*		CHARGE, GBPG, ICBP	Endothelial function?	
12q24.2	*TBX5-TBX3*		CHARGE, ICBP		
15q24.1	*CSK*		CHARGE, GBPG, AGEN-BP, ICBP	Vascular function	
16p12.3	*UMOD*		BP-extremes	Renal electrolyte balance? Renal function?	
17q21.3	*ZNF652*		GBPG, AGEN-BP, ICBP		
20q13	*GNAS-EDN3*		GBPG, ICBP	Vascular function	

Source: Adapted from Padmanabhan et al. [202].

Note: Key genes at each locus are shown with their known or potential BP-regulating pathway and established antihypertensive treatment. GWAS studies: AGEN [7], BP-Extremes [10], CHARGE [8], GBPG [9], Gene-centric [6], HYPERGENES [11], and ICBP [5].

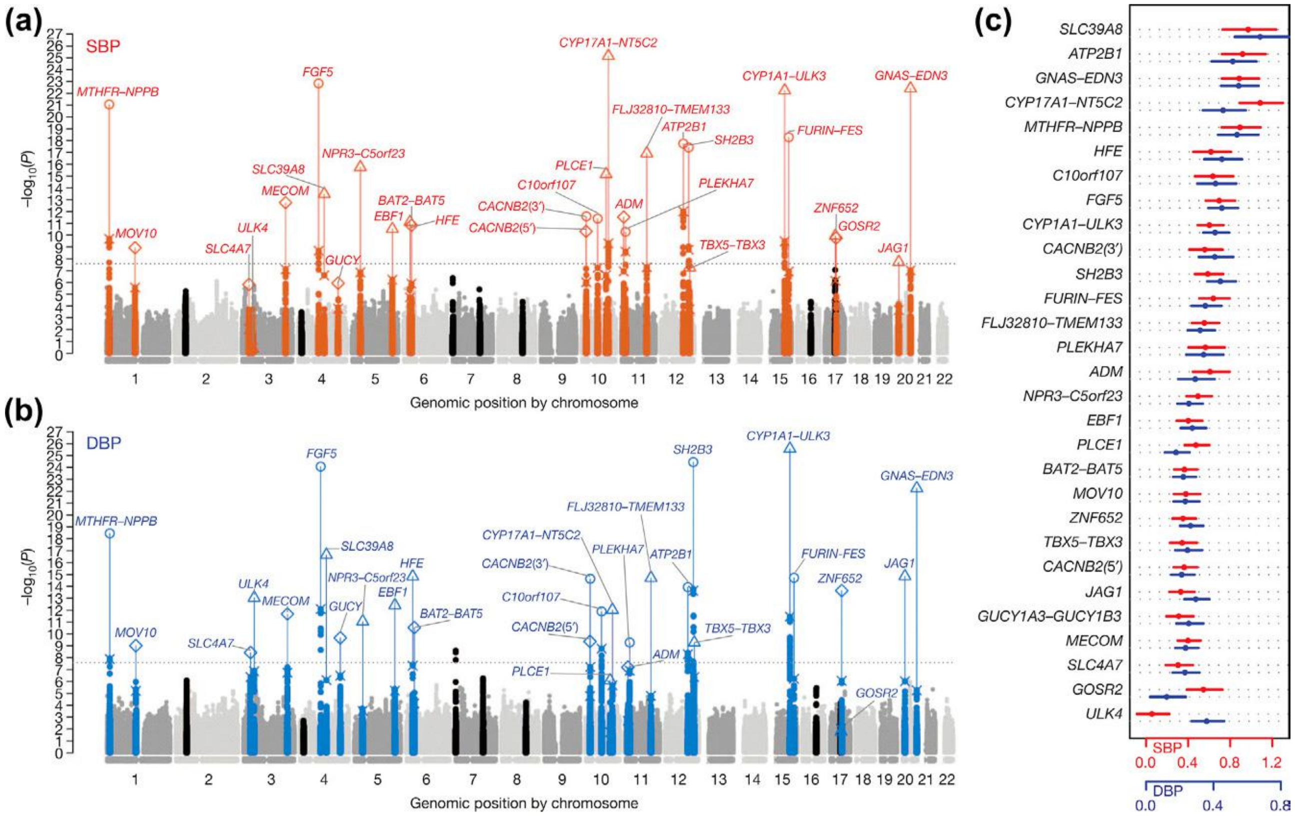


FIGURE 32.1 Genome-wide $-\log_{10}$ p-value plots and effects for significant loci. The plots are shown for systolic (SBP: panel a) and diastolic (DBP: panel b). SNPs in loci reaching genome-wide significance are labeled in red for SBP and blue for DBP (±2.5 Mb of lowest p-value); lowest p-values in the initial genome-wide analysis as well as the results of analysis, including validation data, are labeled separately. The lowest p-values in the initial GWAS are denoted by "X." The range of sample sizes in the final meta-analysis, including the validation data, are denoted by a circle (96–140 k), triangle (>140–180 k), and diamond (>180–220 k). SNPs near unconfirmed loci are in black. The horizontal dotted line is $p = 2.5 \times 10^{-8}$. Panel c shows the effect size estimates and 95% confidence bars per BP-increasing allele of the 29 significant variants for SBP (red) and DBP (blue). Effect sizes are expressed in mmHg/allele. GUCY—*GUCY1A3-GUCY1B3*. *Source: From Ehret, et al.* [47] *with permission.*

the adrenocorticotropic hormone (ACTH). It results from a chimeric gene containing the 5′ regulatory sequences of 11β-hydroxylase (*CYP11B1*) (which confers ACTH responsiveness) fused with the distal coding sequences of aldosterone synthase (*CYP11B2*), leading to the use of ACTH rather than angiotensin II or potassium as the main controller of aldosterone secretion [37]. In contrast, familial hyperaldosteronism type 2 (FH-II, OMIM #605635) is an autosomal dominant syndrome due to hyperplasia or adenoma of the aldosterone-producing adrenal cortex. The genetic abnormality causing FH-II has been localized to chromosome 7p22 [38], though the causative gene has not yet been identified. Apparent mineralocorticoid excess (OMIM #218030) is another low-renin hypertension syndrome accompanied by hypokalemia and metabolic alkalosis. The main defect in apparent mineralocorticoid excess is absence or reduced activity of 11β-hydroxysteroid dehydrogenase (*HSD11B2*), resulting in hypertension in which cortisol acts as a potent mineralocorticoid [39]. Finally, defects in enzymes of cortisol biosynthesis result in autosomal recessive disorders collectively called congenital adrenal hyperplasia (CAH) [40]—11β-hydroxylase (OMIM #202010; *CYP11B1*), 3β-hydroxysteroid dehydrogenase (OMIM #613890; *HSD3B2*), 17α-hydroxylase (OMIM #609300; *CYP17A1*), and cholesterol desmolase (OMIM #118485; *CYP11A1*).

Monogenic low-renin hypertension syndromes can also be caused by mutations in renal ion transporters. Pseudohypoaldosteronism type II (Gordon's syndrome; familial hyperkalemia; OMIM #145260) is an autosomal dominant form of hypertension associated with hyperkalemia, nonanion gap metabolic acidosis, and increased salt reabsorption by the kidney. It is caused by mutations in the WNK kinases [41]. Gordon's syndrome thus results from either gain-of-function mutations in *WNK1* or loss-of-function mutations in *WNK4*. Liddle's syndrome (OMIM #177200) is an autosomal dominant condition with a clinical picture of hypertension and aldosterone excess, but with very low aldosterone and renin levels, caused by mutations in the genes coding the beta or gamma subunits of ENaC (*SCNN1B*, *SCNN1G*) resulting in deletions of proline-rich regions [42,43]. These regions facilitate binding of Nedd4-2 (*NEDD4L*), a regulatory repressor that promotes channel degradation. The inability of beta and gamma subunits to bind Nedd4-2 results in constitutive expression of sodium channels and prolongs the half-life of ENaCs at the renal distal tubule apical cell surface, leading to increased rates of sodium reabsorption, volume expansion, and hypertension.

Conversely, mutations that reduce salt retention, such as those associated with Bartter's (*SLC12A1, KCNJ1, CLCNKB, BSND, CaSR, ClCK-A*) syndrome and Gitelman's (*SLC12A3*) syndrome, tend to lower BP and protect against hypertension [36,44]. In addition to the kidney, the heart secretes a family of vasodilatory and natriuretic hormones in response to increased wall stress—atrial natriuretic peptide (*NPPA*) and B-type natriuretic peptide (*NPPB*). Knockout of one copy of *NPPA* in mice increases BP, while overexpression of *NPPA* lowers BP [45,46]. There is now convincing evidence that common variations in the *NPPA-NPPB* locus influence levels of natriuretic peptides and BP in opposite directions [47–50]. Furthermore, a SNP near the natriuretic peptide clearance receptor (*NPR3*) also showed genome-wide significant association in European, African, and Japanese BP GWASs [47,50,51]. A GWAS for hypertension using an extreme case-control design identified a SNP in the 5′ region of the uromodulin gene (*UMOD*), which is almost exclusively expressed in the thick ascending limb of the loop of Henle in the kidney. This SNP identifies a potentially novel pathway of BP regulation through an effect on sodium homeostasis [52]. Moreover, independent studies have identified SNPs that are highly correlated with the hypertension SNP near *UMOD* to be associated with chronic kidney disease [53].

32.3.2 Autonomic Nervous System

The autonomic nervous system maintains cardiovascular homeostasis via pressure, volume, and chemoreceptor signals through three endogenous catecholamines (norepinephrine, epinephrine, and dopamine) synthesized in the adrenal medulla (epinephrine) and cytosol of adrenergic neurons. Adrenergic reflexes modulate BP over the short-term, and adrenergic function, in concert with hormonal and volume-related factors, contributes to the long-term regulation of arterial pressure. Norepinephrine has more affinity for α receptors, whereas epinephrine binds more avidly to β receptors. α1 Receptors (*ADRA1A, ADRA1B, ADRA1D*) are located on postsynaptic cells in smooth muscle and elicit vasoconstriction. α2 Receptors are localized on presynaptic membranes of postganglionic nerve terminals and act as negative feedback controllers, inhibiting further norepinephrine release. Activation of sinus node and myocardial β1 receptors (*ADRB1*) stimulates, respectively, the rate and strength of cardiac contraction and consequently increases cardiac output. β1 Receptor activation also stimulates renin release from the kidney. Activation of β2 receptors (*ADRB2*) by epinephrine relaxes vascular smooth muscle and results in vasodilation.

Pheochromocytomas (PCCs) and paragangliomas (PGLs) are rare neuroendocrine tumors of the adrenal glands and the sympathetic and parasympathetic paraganglia. The annual incidence of PCC is around 2–4 per million. PCCs and sympathetic PGLs are very similar histologically as well as functionally and cause hypertension, which may be either paroxysmal or sustained. About 30% of PCCs and PGLs are currently believed to be caused by germline mutations. Autosomal dominant inherited pheochromocytomas are due to a variety of *RET* proto-oncogene mutations. Other pheochromocytoma susceptibility genes include *RET* (multiple endocrine neoplasia syndrome type 2 (MEN-2)), the tumor suppressor gene *VHL* observed in families with von Hippel–Lindau syndrome, and the gene that encodes succinate dehydrogenase subunit B and D (*SDHB, SDHD*) and causes familial PGL [54]. Incidence and prevalence of malignant PGL are higher in carriers of the SDHB mutation than in carriers of the SDHD mutation, but lower in carriers of the SDHB mutation [55].

32.3.3 Vascular Mechanisms

Vascular diameter and compliance of resistance arteries are important determinants of arterial pressure. Vascular changes in hypertension may result from either abnormal extracellular stimuli or an altered intracellular signaling cascade, leading to enhanced vasoconstriction, blunted vasodilation, and vascular wall hypertrophy/remodeling. All of these contribute to elevated peripheral vascular resistance. Among the humoral factors that regulate arteries in hypertension are: (1) vasoconstrictor agents such as angiotensin II, endothelin-1, catecholamines, and vasopressin; (2) vasodilator agents such as nitric oxide, endothelium-derived hyperpolarizing factor, and natriuretic peptides; (3) growth factors such as insulin-like growth factor-1, platelet-derived growth factor (PDGF), epidermal growth factor (EGF), and basic fibroblast growth factor; and (4) cytokines such as transforming growth factor-β, tumor necrosis factor, and interleukins [56].

Vascular endothelial function also modulates vascular tone. The vascular endothelium synthesizes and releases a spectrum of vasoactive substances, such as nitric oxide, a potent vasodilator. Nitric oxide exerts vasodilating and antiproliferative effects on smooth-muscle cells and inhibits thrombocyte aggregation and leukocyte adhesion. The synthesis of nitric oxide is controlled by the enzyme endothelial nitric oxide synthase (*NOS3*), and is induced by calcium-mobilizing agents and fluid shear stress. Other vascular relaxation factors include endothelins and prostacyclin. Endothelin-1 (*EDN1*) activates specific ETA receptors (*EDNRA*) on vascular smooth-muscle cells to cause vasoconstriction and cell proliferation. In contrast, endothelial ETB receptors (*EDNRB*) mediate vasodilation via release of nitric oxide and prostacyclin (*PGI2*).

Ion transport by vascular smooth-muscle cells may contribute to hypertension-associated abnormalities of vascular tone and vascular growth, both of which are modulated by intracellular pH. An increased Na^+/H^+ exchanger (*SLC9A1*) can stimulate vascular tone and cell growth by increasing sodium reabsorption in renal proximal tubule cells [57]. Other ion transporters considered are Na^+ bicarbonate transporter (*SLC4A10*) and the Na^+-Ca^+ exchanger NCX (*SLC8A1*). One hypothesis for the mechanism by which excess salt intake elevates BP is through the observed rise in cardiotonic steroids such as ouabain in response to salt intake [58]. It is believed that ouabain inhibits the plasma membrane Na^+/K^+ ATPase, leading to an increase in cytosolic Na^+ concentration that raises the cytosolic Ca^{2+} concentration through the involvement of the Na^+/Ca^{2+} exchanger NCX (*SLC8A1*) and thereby increases contraction in vascular or heart muscle [59].

32.4 NEED FOR STRATIFIED APPROACHES TO TREATMENT WITH ANTIHYPERTENSIVES

Hypertension substantially increases risks for stroke, coronary heart disease (including heart attack), renal failure, and heart failure, and it is the most common chronic disease for which medications are prescribed [60]. Clinical trials comparing different antihypertensive treatments with placebo have shown that effective antihypertensive treatment can reduce mortality from stroke by 35–40%, myocardial infarction by 20–25%, and heart failure by >50% [61].

The guidelines on antihypertensive therapy around the world—such as from the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC) in the United States [62], the Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice [63], the Chinese Hypertension Guideline Committee [64], and the Japanese Society of Hypertension [65]—all considered four drug classes as appropriate first-line therapy for hypertension: thiazide diuretics and thiazide-like diuretics; calcium channel blockers (CCBs); angiotensin-converting enzyme (ACE) inhibitors; and angiotensin receptor blockers (ARBs). β-Blockers are no longer considered an appropriate first-line therapy [63]. Despite multiple first-line therapies for hypertension, less than 50% of hypertensive patients worldwide have their BP controlled, with the control rate ~10% in the Middle East, 13% in Northern Europe, 20% in Southern Europe, 24% in Asia, 28% in North America, and 31% in Latin America. This is according to a global study in 2010 [66]. The control rates in the most recent study were higher, but still all below 50% [67].

Additionally, considerable interindividual variability has been documented in BP responses to agents from all classes of antihypertensive drugs [68,69], and response rates to monotherapy with any given antihypertensive drug are only about 50% [70]. Current selection of initial antihypertensive therapy is essentially by trial and error. The difficulty in determining the most appropriate medication for a specific patient likely contributes to the fact that less than half of hypertensive patients worldwide currently have their BP controlled.

Stratifying hypertensive patients based on the predictors of BP response to antihypertensive drugs can potentially guide antihypertensive therapy to improve the control rate and better prevent long-term adverse cardiovascular events. The most appropriate end points for antihypertensive treatment are prevention of long-term cardiovascular outcomes such as death, nonfatal myocardial infarction, nonfatal stroke, heart failure, and renal failure. BP response to antihypertensive treatment is a good surrogate marker for predicting reduction in these long-term clinical outcomes. Large clinical trials in which a difference in BP treatment arms was observed usually also showed a significant difference in cardiovascular outcomes. For example, differences in total cardiovascular events were observed in the Anglo-Scandinavian Cardiac Outcomes Trial–Blood Pressure Lowering Arm (ASCOT-BPLA) by treatment strategy [71]. Differences in heart failure events between treatment strategies were observed in the Antihypertensive and Lipid-Lowering to Prevent Heart Attack Trial (ALLHAT) [72]. On the other hand, in large trials where there was no difference in on-treatment BP between treatment strategies, usually there was no difference in long-term outcomes—for example, the INternational VErapamil SR Trandolapril Study (INVEST) [73] and the Nordic Diltiazem Study (NORDIL) [74].

The predictors of BP response to antihypertensive drugs include nongenetic (demographic and clinic) and genetic factors. The majority of the reports on predictors

TABLE 32.2 Studies with Blood Pressure Response Phenotype and GWAS Data

Full Name	Acronym	Mean Pretreatment BP (mmHg)	Mean Age	Antihypertensives (Monotherapy)	Race/Ethnicity (Sample Size)
Genetic Epidemiology of Responses to Antihypertensives [76]	GERA	147/96	49	Hydrochlorothiazide (25 mg)	380 European Americans, 225 African-Americans
Genetic Epidemiology of Responses to Antihypertensives 2 [86]	GERA 2	133/85	49	Candesartan	236 Caucasians, 203 African-Americans
Genetics of Drug Responsiveness in Essential Hypertension Study [77]	GENRES	151/100	51	Amlodipine (5 mg), bisoprolol (5 mg), hydrochlorothiazide (25 mg), losartan (50 mg)	206 Finnish men
Pharmacogenomics Evaluation of Antihypertensive Responses [75]	PEAR	152/98	49	Atenolol (100 mg), hydrochlorothiazide (25 mg)	461 Caucasians, 207 African-Americans
Pharmacogenomics Evaluation of Antihypertensive Responses 2[a]	PEAR 2	146/93	50	Metoprolol (100 mg), chlorthalidone (25 mg)	∼180 Caucasians, ∼120 African-Americans ∼120 Caucasians, ∼80 African-Americans
Nordic Diltiazem [74]	NORDIL	171/102	61	Thiazide diuretics[b]	420 Caucasians
Milan Italian study [98]	Milan study	148/98	46	Hydrochlorothiazide (HCTZ)	215 Caucasians

[a] PEAR2 is ongoing; sample sizes were estimated based on patients enrolled by August 2013. [b] NORDIL had a washout period before treatment and therefore is considered to have valid baseline BP data; it had patients on monotherapy of other antihypertensives, such as diltiazem (CCB) and β-blockers, but the number of patients with the genotype was not published.

of BP response were generated from four clinical trials: the Pharmacogenomics Evaluation of Antihypertensive Responses (PEAR) study [75], the Genetic Epidemiology of Responses to Antihypertensives (GERA and GERA2) [76], and the Genetics of Drug Responsiveness in Essential Hypertension Study (GENRES) [77]. Table 32.2 lists the studies with appropriately measured BP response phenotype after antihypertensive treatment and DNA available for pharmacogenomic association analysis.

32.4.1 Nongenetic Approaches

The nongenetic predictors of BP response include pretreatment BP, age, race, and plasma renin activity, with pretreatment BP and plasma renin activity being the most consistent. The effect of pretreatment BP is nonspecific to antihypertensive drug class, while age, race, and plasma renin activity are specific to antihypertensive drug class. Other factors have been shown to be associated with BP response to certain antihypertensives in some studies but not in others.

32.4.1.1 *Pretreatment Blood Pressure*

The GERA study first reported in 2002 that higher baseline BP was associated with better response to four weeks of treatment with hydrochlorothiazide (HCTZ, a thiazide diuretic) [76]. An Italian study showed that baseline SBP, age, and gender were associated with SBP response after treatment with the ACE inhibitor fosinopril [78]. The PEAR study showed that pretreatment BP is a significant predictor of SBP and DBP response to atenolol (β-blocker) and HCTZ both as monotherapy and as add-on therapy. The higher the pretreatment BP, the greater the reduction achieved by antihypertensive treatments [79]. GENRES reported that placebo BP level was the best predictor for BP responses to four antihypertensive drug monotherapies: amlodipine (CCB), bisoprolol (β-blocker), HCTZ, and losartan (ARB) [80]. Since the association of pretreatment BP with BP response was observed in all five antihypertensive drug classes, it appears to be nonspecific to drug class.

32.4.1.2 *Plasma Renin Activity*

The RAAS plays an important role in BP regulation, and hypertensive patients are characterized by interindividual variation in its activity [81]. In the early 1970s, Laragh and colleagues first proposed measurement of plasma renin activity as a method of characterizing relative participation of a patient's vasoconstriction and volume components [82,83]. They suggested that low activity corresponds to better BP response to diuretics and CCBs, while high activity favors a response to ACE inhibitors and β-blockers [84]. This theory has been confirmed by several studies.

In the Veterans Affairs Cooperative Study on Single-Drug Therapy of Hypertension trial, a total of 1105 ambulatory men were randomized to 1 of 6 antihypertensive drugs: HCTZ, atenolol, captopril (an ACE inhibitor), clonidine (an α2 receptor blocker), diltiazem (CCB), or prazosin (an α1 receptor blocker). The BP control rate was evaluated by comparing the percentage of patients achieving a goal of <140/90 mmHg. It was found that HCTZ was best in low- and medium-renin profiles and that captopril was best in medium- and high-renin profiles. Clonidine and diltiazem had consistent response rates regardless of renin profile [85]. This study concluded that in men with stage 1 and stage 2 hypertension, although BP responses to different antihypertensives were consistent with baseline renin profile, age and race were better predictors of BP response.

The GERA study showed that lower baseline plasma renin activity was associated with significantly greater BP response to HCTZ [76]. Later, investigators reported that higher baseline plasma renin activity predicted greater response to candesartan (ARB) [86]. Their findings supported the notion that characteristics associated with poor response to diuretic therapy are associated with better ARB response [76,86]. The GENRES study found that, in Finnish hypertensive men, plasma renin activity was positively correlated with BP responses to losartan (ARB) and bisoprolol (β-blocker), and negatively correlated with BP response to HCTZ and amlodipine (CCB) [87].

Analysis of the PEAR study showed that plasma renin activity was an independent predictor of SBP and DBP responses to atenolol and HCTZ, not only in monotherapy but also in add-on therapy. The predictive effects of plasma renin activity were directionally opposite for atenolol and HCTZ responses, with higher renin predicting greater BP responses to atenolol as monotherapy or as an add-on and lesser responses to HCTZ as monotherapy or as an add-on (Figure 32.2).

A recent retrospective analysis of the GERA study assessed the BP control rates of three different treatment strategies among patients with stage 1 essential hypertension: (1) Thiazide diuretic (HCTZ) for all; (2) thiazide diuretic for all black subjects and white subjects age ≥50 years, and a RAAS blocker (candesartan) for white subjects aged <50 years; (3) thiazide diuretic (HCTZ) for plasma renin activity <0.6 ng/ml/h (suppressed renin activity) and a renin–angiotensin system blocker (candesartan) for plasma renin activity ≥0.6 ng/ml/h (nonsuppressed renin activity). It was found that the renin-based strategy was associated with the highest control rate of 69.4% versus 61.3% for the age/race strategy ($p<0.0001$) and 53.8% for the thiazide for all strategies ($p<0.0001$). This study also found that not all of the racial difference in response to candesartan versus HCTZ can be explained by differences in plasma renin activity status [88].

32.4.1.3 *Age and Race*

Age and race, as predictors of BP response to antihypertensives, are correlated to plasma renin activity. Older age and black race are used as markers for low activity, whereas younger age and white race are markers for higher activity. The British Hypertension Society recommends the use of a treatment algorithm based on age and race [89]. Hypertension can be broadly classified as "high renin" or "low renin" and is therefore best treated initially with one of two categories of antihypertensive drugs: those that inhibit the renin–angiotensin system—ACE inhibitors and agiotensin receptor blockers (A drugs) and β-blockers (B drugs); and those that do not inhibit it—calcium channel blockers (C drugs) and diuretics (D drugs). People who are younger than 55 and white tend to have higher renin concentrations than those aged 55 or older or who are black (of African descent). A or B drugs are therefore generally more effective as initial BP-lowering treatments in younger white patients than are C or D drugs [89].

GENRES reported that in moderately hypertensive Finnish men, age was positively associated with BP responses to amlodipine and HCTZ [77]. In the GERA study, black race was associated with significantly greater SBP and DBP responses to HCTZ, and older age was a predictor of better DBP, but not SBP response, to HCTZ [76].

The PEAR study has confirmed the racial differences in BP response to HCTZ versus atenolol. In this study, where uncomplicated hypertensive patients were randomized to either HCTZ or atenolol monotherapy, African-Americans showed a greater BP-lowering effect than whites with HCTZ (−13.0/−7.4 mmHg versus −8.0/−4.2 mmHg, $p<0.001$), but a smaller BP-lowering effect than whites with atenolol monotherapy (−1.1/−2.9 mmHg versus −9.9/−9.2 mmHg, $p<0.0001$) [90].

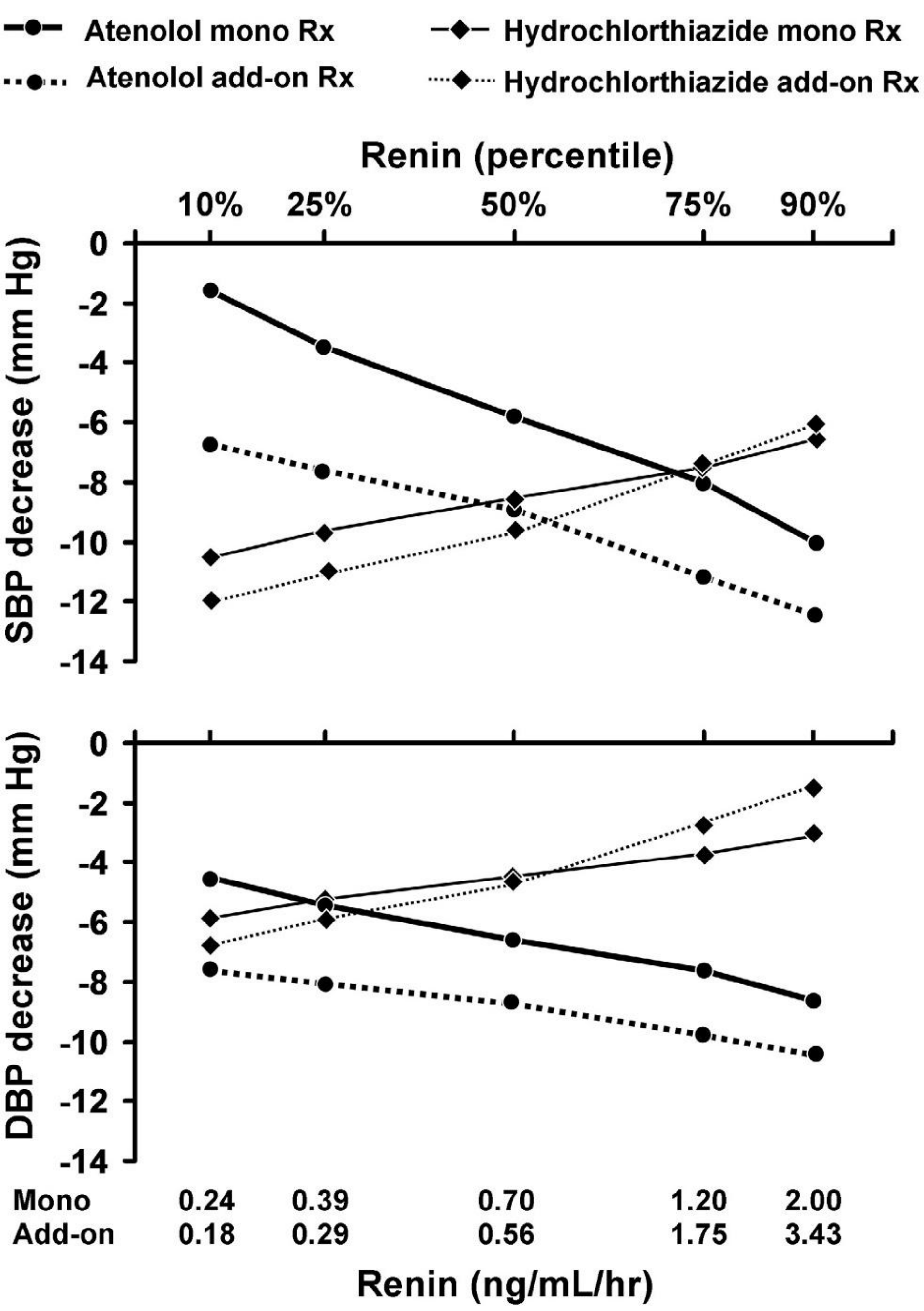


FIGURE 32.2 **Predicted blood pressure responses as a function of plasma renin activity.** The plotted points are at the observed 10th, 25th, 50th, 75th, and 90th percentiles of plasma renin activity distributions prior to mono- and add-on therapy. Participants were randomized to receive atenolol monotherapy followed by hydrochlorothiazide add-on therapy ($N = 180$) or hydrochlorothiazide monotherapy followed by atenolol add-on therapy ($N = 183$). The predicted blood pressure responses at each renin percentile are from a single linear regression model in which all other predictors (race, sex, age, log hypertension duration, current smoker, ethanol intake, height, waist circumference, and baseline blood pressure) were held constant at their respective mean values. DBP—diastolic blood pressure; SBP—systolic blood pressure. *Source: From Turner et al.* [79] *with permission.*

32.4.1.4 Other Factors

In a study of more than 3000 untreated Chinese hypertensive individuals randomized to eight weeks of antihypertensive monotherapy, women were shown to have better SBP reduction from atenolol than men, with a mean difference of 3.9 mmHg. It was also shown that women had better reduction in DBP after HCTZ treatment compared to men, with a mean difference of 1.8 mmHg. However, there were no gender differences in BP response to nifedipine-SR (a CCB) and captopril [91]. The PEAR study also showed that females responded better to atenolol in SBP and DBP as monotherapy (−4.21/−0.79 mmHg) and as an add-on (−3.73/−2.39 mmHg), while no evidence of gender differences was found in response to HCTZ [79]. However, the GERA study showed that females responded better to HCTZ for both SBP and DBP [76].

Other factors that did not show consistent association with BP response to antihypertensives include duration of hypertension [79], current smoker (atenolol add-on therapy) [79], and alcohol intake (SBP response to HCTZ add-on therapy) [79]. GENRES also reported that serum total calcium concentration and daily urinary excretion of sodium were both negatively correlated with BP response to amlodipine, but not in response to the other three drugs [87].

In summary, the most consistent nongenetic predictors of BP response to antihypertensives are pretreatment BP and plasma renin activity, while demographical factors such as age and race can be used to capture information for plasma renin activity.

32.4.2 Genetic Approaches

Several easily determined clinical characteristics have been shown to predict BP responses to commonly prescribed antihypertensive drugs, but the majority of interindividual BP response variation remains unexplained. The GERA study reported that the combined effects of all identified predictors accounted for 38% of interindividual variation in SBP response and 20% in DBP response to HCTZ [76]. The predictors of BP response to candesartan accounted for 39% and 33% of interindividual variation in SBP and DBP responses, respectively [86]. The PEAR study reported that the percentage of variation in BP response explained by baseline BP, age, race, and plasma renin activity was 16% for DBP and 21% for SBP response after HCTZ monotherapy, and 27% and 29%, respectively, after atenolol monotherapy [79].

Pharmacogenomics, which utilizes genetic information to determine drug response, is one tool that might be useful in further individualizing antihypertensive drug selection. We discuss the pharmacogenomics of BP response in Section 32.5 and long-term cardiovascular outcomes after antihypertensive therapy in Section 32.6. Our focus is on genetic variants that meet one of the following criteria: (1) replication in an independent sample or a different race/ethnicity group; (2) association with multiple related phenotypes, such as prior association with BP or hypertension; and (3) functional basis. We do not discuss a single association reported in the literature or ones for which there were problems that we believe bring the data into question (such as deviation from Hard-Weinberg equilibrium or inappropriate study design).

32.5 PHARMACOGENOMICS OF BLOOD PRESSURE RESPONSE TO ANTIHYPERTENSIVE DRUGS

Most pharmacogenomics studies of antihypertensives have used the candidate gene approach, which focuses on genetic variations that can influence antihypertensive drug pharmacokinetic or pharmacodynamic mechanisms. The first GWAS of an antihypertensive drug was published in 2008 by Turner and colleagues and was based on an analysis of BP response to HCTZ in the GERA study [92]. Since then two more hypertension pharmacogenomic GWAS have been published, one on BP response to HCTZ and the other on BP response to the ARB candesartan. As of June 2013, there was no GWAS published on BP response to β-blockers, CCBs, and ACE inhibitors.

The genes with the most compelling evidence of antihypertensive pharmacogenomic association with BP response have been those with a prior BP or hypertension association or those with functional evidence of involvement with BP regulation and replication across different populations, either in independent studies or in different race/ethnicity groups. The top genes from the candidate genes studies are *NEDD4L*, which is associated with BP response to thiazide diuretics, and *ADRB1*, which is associated with BP response to β-blockers. The most compelling genes from GWAS analyses are SNPs in the chromosome 12 region associated with HCTZ response in African-American patients, and those in the *PRKCA* and *GNAS-EDN3* regions associated with HCTZ response in European patients.

32.5.1 Thiazide Diuretics

Multiple hypertension treatment guidelines from different regions of the world recommend thiazide or thiazide-like diuretics as a first-line treatment for patients with essential hypertension [62–65]. In most countries, thiazide diuretics are the cheapest antihypertensives available and the mainstay of antihypertensive pharmacotherapy. HCTZ is the second most commonly prescribed antihypertensive in the United States, with approximately 135 million prescriptions dispensed in 2008 either in monotherapy or in combination therapy [93]. Despite the widespread use of this drug class, the mechanism by which thiazide diuretics lower BP remains poorly understood. They achieve their diuretic effect by decreasing sodium reabsorption via inhibition of the Na^+/Cl^- cotransporter in the renal distal convoluted tubule [94].

Thiazide use acutely results in an increase in fluid loss to urine and a decrease in extracellular fluid and plasma volume, which results in diminished venous return, increased renin release, reduced cardiac output, and decreased BP [94]. However, this does not explain the chronic BP-lowering effect of thiazides, because the plasma and extracellular fluid volumes almost fully recover 4–6 weeks after thiazide initiation, although BP reduction is maintained. Possible mechanisms for chronic BP lowering include direct vasodilation mediated by endothelial or vascular smooth muscle and indirect compensation for acute decreases in cardiac output [94]. There have been many candidate genes studies on the pharmacogenetic association of BP response and thiazide diuretics, most of which used HCTZ.

32.5.1.1 ADD1

Alpha adducin, *ADD1*, is a ubiquitously expressed protein that is part of the cytoskeleton and may modulate ion transport. In a family linkage study in 1997, the nonsynonymous SNP Gly460Trp in the *ADD1* gene that encodes α-adducin was found to be associated with sensitivity to changes in sodium balance. Patients who were carriers of the 460Try allele had a lower baseline plasma renin activity and a greater BP-lowering response to two months of HCTZ treatment compared to Gly/Gly homozygotes (mean arterial BP decrease of 15.9 versus 7.4 mmHg, $p = 0.001$) [95]. A later study, utilizing erythrocytes as a mirror to transport in the cell membrane, reported that the Gly460Trp polymorphism may affect renal Na handling through an alteration in ion transport across the cell membrane [96]. The association between this polymorphism and BP response to thiazide diuretics has been confirmed by some subsequent studies [97,98] but not others [99–102].

32.5.1.2 GNB3

The T allele of the C825T polymorphism of the gene encoding the G-protein β3 subunit, *GNB3*, has been associated with low plasma renin and elevated DBP [103]. This suggests that the T allele might also be associated with a volume-expanded, sodium-sensitive, and therefore diuretic-responsive form of hypertension. Turner and colleagues tested whether this polymorphism was associated with BP response to HCTZ in the GERA study of 197 blacks and 190 non-Hispanic whites. They found that the T allele was associated with better response to HCTZ with a gene-dose effect, even after adjusting for other significant predictors [104]. This association was later confirmed by an independent study in the Netherlands [105] but not replicated by others [101].

32.5.1.3 NEDD4L

To date, the candidate gene with the most compelling evidence for thiazide pharmacogenomic association is *NEDD4L*, the neural precursor cell expressed developmentally down-regulated 4-like gene (also known as *NEDD4-2*), encodes a ubiquitin ligase that plays a critical role in epithelial sodium transport by regulating the cell surface expression of the epithelial sodium channel (ENaC). By targeting ENaC for degradation, *NEDD4L* is a significant determinant of sodium reabsorption in the distal nephron. Several studies have shown that polymorphisms in *NEDD4L*—specifically rs4149601 G>A, which is known to result in a cryptic splice site—are associated with hypertension in multiple populations [106,107]. In a prior human study, G carriers of this SNP were also shown to have higher salt sensitivity and lower plasma renin concentration compared to the AA genotype [108].

Based on these data, one would expect individuals who carry the G allele of rs4149601 to respond better to thiazide diuretics. This hypothesis was confirmed in the NORDIL study, where patients treated with a combination of thiazide diuretics and β-blockers, or either drug alone, were compared to patients treated with the CCB diltiazem. In this study, rs4149601 G allele carriers were shown to have greater BP response than patients with the AA genotype (SBP: −19.5 versus −15.0 mmHg; DBP: −15.4 versus −14.1 mmHg) when treated with a thiazide diuretic/β-blocker. No differences in response to diltiazem by genotype were shown [109].

The PEAR study evaluated the association of the G allele of rs4149601 with BP response to HCTZ and atenolol separately, finding a significant association in HCTZ-treated patients (Figure 32.3), where there was

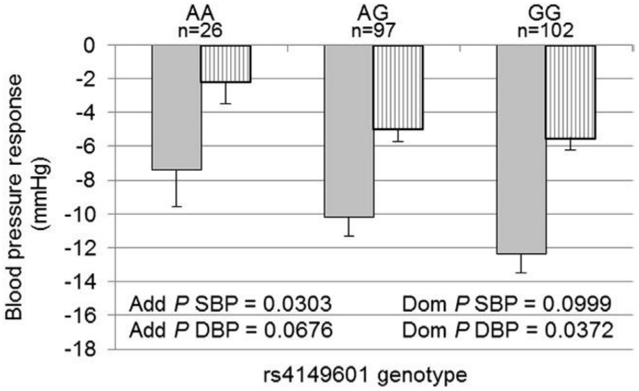


FIGURE 32.3 **Blood pressure response by NEDD4L rs4149601 genotype in white patients treated with HCTZ in PEAR.** Solid gray bars indicate ΔSBP; gray and white lined bars indicate ΔDBP. Values are shown as means ± standard error. Add—additive, Dom—dominant. *Source: From McDonough et al.* [110] *with permission.*

better BP reduction when the G allele was increased (SBP: −12.4 mmHg, −10.2 mmHg, and −7.4 mmHg; DBP: −5.5 mmHg, −5.0 mmHg, and −2.2 mmHg for GG, GA, and AA). However, there was no evidence of such association in the atenolol-treated patients [110]. The G allele was also shown to be associated with cardiovascular outcome in the NORDIL and INVEST studies (see the next section).

32.5.1.4 GWAS Signals

Three GWASs of hypertensive pharmacogenomics had been published as of June 2013. Two of them reported pharmacogenomic associations between BP response and HCTZ treatment, one in African-Americans and the other in European Americans.

The first GWAS analysis of hypertension pharmacogenomics was published in 2008 and used the GERA study as a basis. A haplotype was found to be associated with DBP response to HCTZ in African-Americans [92], located on chromosome 12q15 (consisting of SNPs rs317689, rs315135, and rs7297610 near *LYZ*, *YEATS4*, and *FRS2*, of which rs7297610 is the strongest). This association was later replicated in an independent sample of PEAR African-American hypertensive patients treated with HCTZ as monotherapy or as an add-on [111]. It was not observed in patients treated with atenolol (Figure 32.4).

The PEAR study also explored expression of the genes *YEATS4* and *FRS2* pre- and post-HCTZ treatment, finding that the CC genotype (associated with better BP response to HCTZ) had higher baseline expression of the *FRS2* gene compared to the CT genotype, but also finding that there was no difference in the expression level of the *YEATS4* gene [111].

The most recent antihypertensive pharmacogenomic GWAS involved five independent studies, two as discovery cohorts (PEAR and GERA) and three (GENRES, NORDIL, and Milan) as replication cohorts [100]. An intronic polymorphism, rs16960228, in protein kinase C alpha (*PRKCA*) on chromosome 17q24.3 was replicated for the same direction of association with DBP response in the NORDIL and GENRES studies, and the combined four-study meta-analysis p-value achieved genome-wide significance ($p = 3.3 \times 10^{-8}$). SBP/DBP responses were consistently greater in rs16960228 A allele carriers than in GG homozygotes (>4/4 mmHg) across study samples (Figure 32.5).

This association was further validated in three ways. First, in PEAR rs16960228 was also associated with DBP response to the β-blocker, compared with the opposite-direction-of-association compared response to HCTZ (which was expected given the different pharmacologies of these drug classes). Second, SNPs in *PRKCA* were significantly associated with DBP response in PEAR African-Americans. Third, gene expression analysis using RNA demonstrated that the carriers of the rs16960228 variant A allele had significantly greater relative *PRKCA* expression level than GG homozygotes [100].

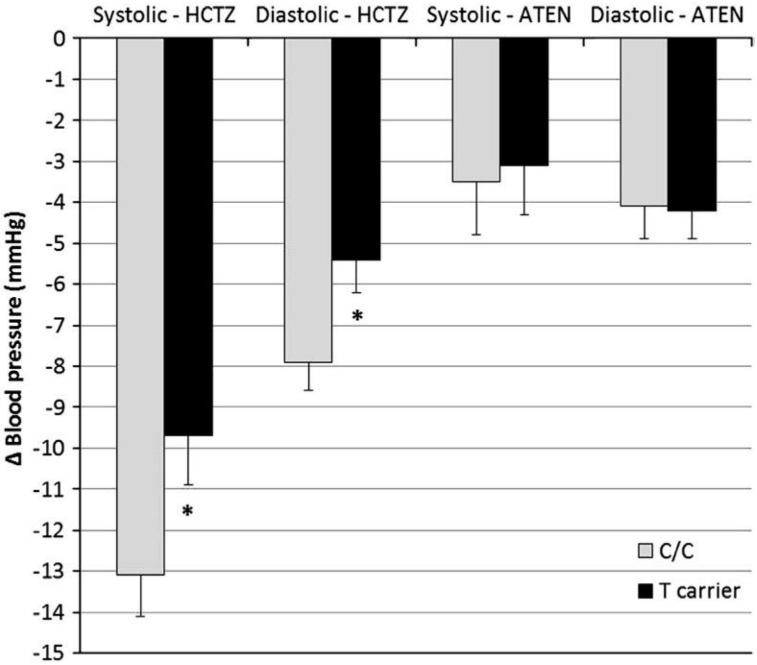


FIGURE 32.4 **Blood pressure response to HCTZ and atenolol by rs7297610 genotype in PEAR African-Americans.** Values are adjusted for age, sex, and baseline blood pressure. Error bars indicate s.e. $*p \leq 0.05$ compared with the common C/C genotype. *Source: From Duarte J, et al.* [111] *with permission.*

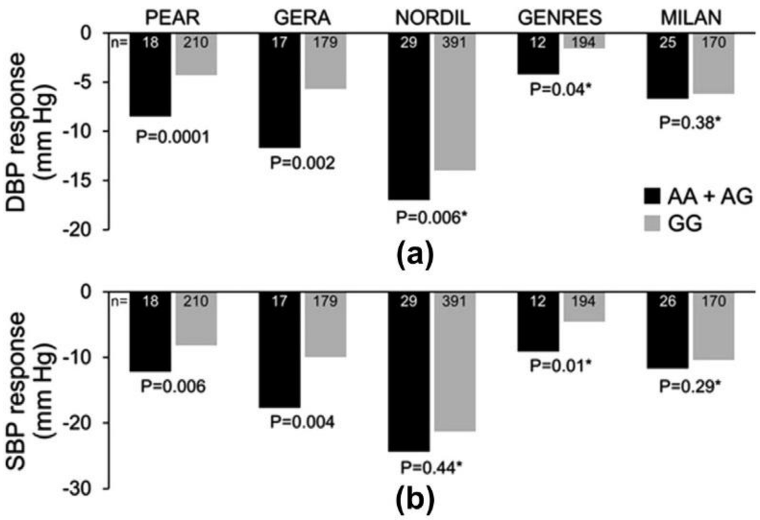


FIGURE 32.5 **Blood pressure response to hydrochlorothiazide by chromosome 17 rs16960228 genotype of participants from five independent studies.** (a) diastolic blood pressure response; (b) systolic blood pressure response. The blood pressure responses are adjusted for pretreatment blood pressure levels, age, and sex; p-values are for contrast of adjusted means between genotype groups. *Source: From Turner et al.* [100] *with permission.*

The rs2273359 polymorphism in the *GNAS-EDN3* region shows a consistently significant association of SBP response to HCTZ in PEAR, GERA, and NORDIL (data for this SNP were not available in GENRES and the Milan study). Among the participants in the three studies, CG heterozygotes demonstrated consistently greater BP response to HCTZ than CC homozygotes (no GG homozygotes were observed) (see Figure 32.6). The combined p-value from the three-study meta-analysis approached genome-wide significance ($p = 5.5 \times 10^{-8}$) [100]. This region was previously associated with BP and hypertension in GWAS analyses.

32.5.2 β-Blockers

The most compelling pharmacogenomics association for BP response to β-blockers is that of *ADRB1*, which was shown to be associated not only with BP response, but also with long-term cardiovascular outcomes.

32.5.2.1 ADRB1

ADRB1 encodes the β1-adrenergic receptor, which is the primary protein target of all marketed β-blockers. It contains two nonsynonymous polymorphisms, one at codon 389, Arg389Gly, and one at codon 49, Ser49Gly. There is substantial evidence for functional consequences for both polymorphisms, but particularly for Arg389Gly. In membrane adenylyl cyclase assays from recombinantly expressing cells, the Arg389 form of *ADRB1* had greater basal and against-stimulated activity compared with the Gly389 form; therefore, it also had greater downstream signaling in response to agonist [112]. The Ser49Gly polymorphism lies in the region that may play a role in membrane insertion, expression, trafficking, and agonist-promoted down-regulation [113].

Two independent studies found enhanced agonist-promoted down regulation with the Gly49 form of the receptor [114,115]. Based on functional data for these two SNPs, one would hypothesize that the more responsive Ser49 and Arg389 forms have the greatest β-blockade efficacy. Consistent with this hypothesis, several studies have shown that patients with the Arg389Arg genotype or those with the Ser49Arg389 haplotype have the greatest BP response to β-blockers (metoprolol) [116–118].

Despite some studies that failed to show such an association [119,120], this gene is of particular interest because Gly389 has also been associated with decreased SBP and DBP in a large genetic study evaluating 30 regions that code for known antihypertensive drug targets [121]. More important, this polymorphism has also been associated with cardiovascular outcomes after antihypertensive treatment in a large clinical trial (see the next section).

32.5.2.2 GRK4

G protein-coupled receptor kinases are a family of serine/threonine kinases that phosphorylate activated G protein-coupled receptors (GPCRs), leading to receptor desensitization, deactivation, and endocytosis. *GRK4* plays a role in BP homeostasis through phosphorylation of GPCRs that are important to BP regulation, including dopamine receptors and potentially β1-adrenergic receptors—the key protein target of β-blockers. The African American Study of Kidney Disease and Hypertension (AASK) trial evaluated

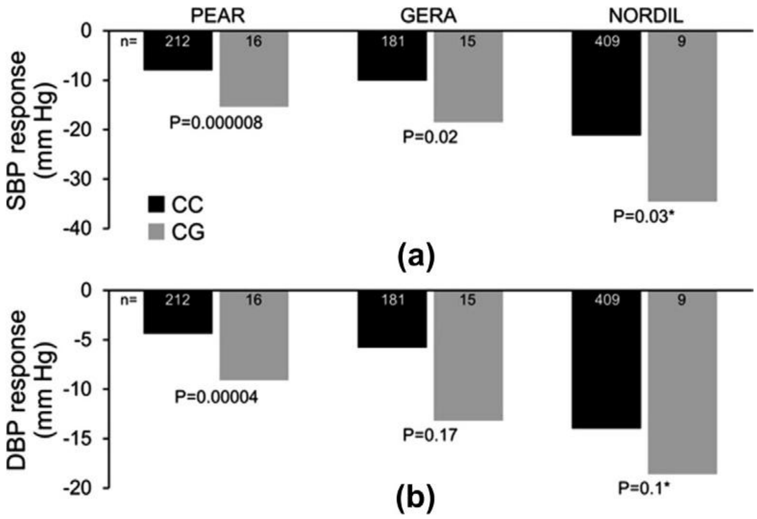


FIGURE 32.6 Blood pressure response to hydrochlorothiazide by chromosome 20 rs2273359 genotype of participants from three independent studies. (a) systolic blood pressure response; (b) diastolic blood pressure response. The blood pressure responses are adjusted for pretreatment blood pressure levels, age, and sex; *p*-values are for contrast of adjusted means between genotype groups. *Source: From Turner et al.* [100] *with permission.*

the association of three nonsynonymous polymorphisms in *GRK4* Arg65Leu (rs2960306), Ala142Val (rs1024323), and Ala486Val (rs1801058). It also evaluated the time taken to reach a mean arterial pressure of 107 mmHg after metoprolol treatment. A significant sex-specific association for Ala142Val was found among African-American men with early hypertensive nephrosclerosis: men with the 142Ala polymorphism were less responsive to metoprolol if they also had a 65Leu variant; this association was not seen in women [122].

Vandell and colleagues also studied the association of these three *GRK4* SNPs with BP response to atenolol in the PEAR study. They found that the variant alleles of all three SNPs showed, or trended toward, a lesser BP lowering. They further evaluated the haplotype of the two more common variants and found that increasing copies of the variant 65Leu-142Val haplotype were associated with significantly reduced atenolol-induced DBP lowering (−9.1 ± 6.8 versus −6.8 ± 7.1 versus −5.3 ± 6.4 mm Hg in participants with, respectively, 0, 1, and 2 copies; p = 0.0088). There was no evidence of such association with HCTZ-induced BP response [123]. The variant alleles of all three *GRK4* SNPs were also associated with increased risk for adverse cardiovascular outcomes [123].

32.5.3 Calcium Channel Blockers

Most candidate genes in the CCB pharmacogenomic literature are those that encode proteins relevant to calcium regulation, such as the L-type calcium channel subunits α1C (*CACNA1C*) and β (*CACNB2*) and the β1 subunit of the calcium-activated potassium channel, also known as the BK channel (*KCNMB1*). The majority of the pharmacogenetic studies with CCBs were based on analysis of verapamil SR (a nondihydropyridine CCB) in the INVEST genetic substudy (INVEST-GENES). While there was no evidence of association of *CACNA1C* and *CACNB2* SNPs and BP responses to verapamil-SR, there were significant differences in long-term cardiovascular outcome by genotype for them [124,125].

32.5.3.1 *KCNMB1*

KCNMB1 encodes the β1 subunit of the large-conductance calcium and voltage-dependent potassium (BK) channel in vascular smooth muscle. *KCNMB1* knockout mice have decreased calcium sensitivity, elevated BP, and cardiac hypertrophy [126,127]. *KCNMB1* has two common nonsynonymous polymorphisms: Glu65Lys (rs11739136) and Val110Leu (rs2301149). The Lys65 variant of the Glu65Lys SNP was associated with decreased risk of diastolic hypertension in a Spanish population and was found to result in a BK channel with enhanced calcium sensitivity. This finding supports more efficient negative feedback on the L-type calcium channels resulting in vascular smooth-muscle relaxation [128]. Additionally, Lyn65 carriers appeared to have significantly lower 5-year risk of cardiovascular events compared with the Gly65 homozygotes [129].

In a study of healthy twins, Val110Leu was one of four polymorphisms associated with heart rate variability and baroreflex activity [130]. INVEST assessed the association of these two nonsynonymous SNPs with BP response in hypertensive patients with coronary artery disease after verapamil-SR monotherapy. The study reported that the SBP response did not differ by *KCNMB1* genotype. However, Lys65 variant carriers achieved BP control earlier than

Glu65Glu individuals (1.47 versus 2.83 months, $p=0.01$) and were less likely to require multiple drugs at the time of BP control (adjusted (OR; 95% CI = 0.43 (0.19–0.95)) [131]. This finding was consistent with a prior report indicating that Lys65 is protective.

32.5.4 ACE Inhibitors and ARBs

ACE inhibitors and ARBs are both inhibitors of RAAS. ACE inhibitors act to prevent the conversion of angiotensin I to angiotensin II and prevent the degradation of bradykinin, while ARBs block angiotensin II receptors.

The candidate genes for pharmacogenomic associations in BP response to these two drug classes have been in those in the RAAS: angiotensin-converting enzyme (*ACE*), angiotensinogen (*AGT*), angiotensin receptor 1 (*AGTR1*), and renin (*REN*). However, the prior associations of *ACE*, *AGT*, and *AGTR1* either have been conflicting or have shown no association. In the case of *REN*, only one study reported an association result.

32.5.4.1 ACE

The *ACE* gene encodes the angiotensin-converting enzyme involved in catalyzing the conversion of angiotensin I into a physiologically active peptide, angiotensin II. Angiotensin II is a potent vasopressor and aldosterone-stimulating peptide that controls BP and fluid-electrolyte balance. This enzyme plays a key role in the RAAS. Strong evidence exists for an association between the ACE insertion and deletion (I/D) polymorphism of intron 16 and plasma ACE activity, with increased levels among those with the D allele 132. Therefore, a majority of pharmacogenetic candidate gene studies of ACE inhibitors and ARBs evaluated the *ACE* I/D polymorphism. However, some showed no association [78,133–140] while others reported conflicting results, with either the I allele or the D allele having better response to ACE inhibitors [141–145].

32.5.4.2 AGT

AGT encodes angiotensinogen, preangiotensinogen, or the angiotensinogen precursor, which is expressed in the liver and is cleaved by the enzyme renin in response to lowered BP. The resulting product, angiotensin I, is then cleaved by ACE to generate the physiologically active enzyme angiotensin II. The protein is involved in maintaining BP and in the pathogenesis of essential hypertension and preeclampsia. Mutations in this gene are associated with susceptibility to essential hypertension, and can cause renal tubular dysgenesis, a severe disorder of renal tubular development.

AGT is another gene often evaluated in pharmacogenetic studies of ACE inhibitors and ARBs, especially the SNP rs699, which leads to the substitution of a threonine (Thr) for a methionine (Met) at codon 235. This SNP has been shown to be functional, with higher plasma AGT levels observed in individuals with the Thr/Thr genotype compared to those with the Met/Met genotype [146]. In studies that found a significant association, it was found that Thr/Thr individuals had better BP response to ACE inhibition than Met/Thr and Met/Met individuals [147,148]. In other studies, no association was found [135,138–140].

32.5.4.3 AGTR1

Angiotensin II is a potent vasopressor hormone and a primary regulator of aldosterone secretion. It is an important effector, controlling BP and volume in the cardiovascular system. *AGTR1* encodes the type 1 receptor that is thought to mediate the major cardiovascular effects of angiotensin II. This gene may play a role in the generation of reperfusion arrhythmias following restoration of blood flow to ischemic or infarcted myocardium. The main polymorphism studied was A1166C, and most studies showed lack of association with BP response to RAAS blockade [78,138–140,149,150].

32.5.4.4 GWAS Findings

The only GWAS on BP response to ARBs to date was published in 2012 on candesartan in the GERA study. Almost 200 European Americans and African-Americans with primary hypertension from the two extremes of BP response after candesartan monotherapy were included in the analysis [151]. SNPs associated with candesartan in European Americans were validated by opposite direction of associations with HCTZ response in a separate European American population and same direction of association in African-Americans. A SNP (rs11020821) on chromosome 11q21 near *FUT4* was associated with BP response to candesartan; another SNP in this region (rs3758785 in *GPR83*) had an opposite direction of association with HCTZ response and the same direction of association with candesartan response in African-Americans. Another interesting locus was rs11649240 on chromosome 16, in amiloride-sensitive sodium channel subunit *SCNN1G*, which is involved in mediating renal sodium reabsorption and maintaining BP when the RAAS is inhibited by candesartan [151].

In summary, most BP response pharmacogenomic associations were published on thiazide diuretics, specifically HCTZ. The candidate genes with the most compelling evidence for pharmacogenomic association are *NEDD4L*, which is associated with BP response to thiazide diuretics, and *ADRB1*, which is associated with BP response to β-blockers. The most compelling genes from GWAS analyses were chromosome 12-region SNPs associated with HCTZ response in African-Americans, and the *PRKCA* and *GNAS-EDN3* regions associated with HCTZ response in hypertensive patients of European descent.

32.6 PHARMACOGENOMICS OF ANTIHYPERTENSIVE TREATMENT—RELATED OUTCOMES

Along with genetic markers that can affect BP response to antihypertensive drugs, there are also genetic markers that can influence long-term cardiovascular outcomes after antihypertensive treatment. Many of these are candidate genes, and some also have a documented role in BP response or an association with hypertension risk. Others are novel findings that are limited to treatment-related outcomes.

32.6.1 Clinical Trials

The majority of the data on the pharmacogenomics of antihypertensive treatment-related outcomes arise from four large hypertension-outcome clinical trials: INVEST [73], NORDIL [74], ALLHAT [72], and LIFE [152] (Table 32.3). These clinical trials have between two and four randomized drug classes, and have high-quality clinical data. The other studies on antihypertensive treatment-related outcomes are population-based, often having observational data from medical records and/or prescription drug records.

The genes with the strongest evidence of a positive association affecting treatment-related cardiovascular outcomes are those with replication in multiple cohorts, association with other related phenotypes, and/or functional evidence that supports the association. These top genes are *ADRB1*, *NEDD4L*, *CACNB2*, *CACNA1C*, and *FGB*, identified in candidate gene approaches, and then *SIGLEC12*, *A1BG*, and *F5*, identified through a genome-spanning approach. Additionally, there are genes with strong evidence for no association, including *ADD1* and *ACE*.

32.6.2 Genes with Strong Evidence of Association with Treatment-Related Outcomes

32.6.2.1 ADRB1

ADRB1 has prior association with BP and hypertension [121], strong evidence for a functional effect, and association with BP response to β-blocker treatment. It was also evaluated for association with long-term cardiovascular outcomes after antihypertensive therapy. The two nonsynonymous SNPs with functional evidence, Ser49Gly and Arg389Gly, were studied in relation to cardiovascular outcomes in individuals treated with β-blockers in a genetic substudy of the INVEST trial [153]. In INVEST, the Ser49-Arg389 haplotype was associated with risk of the primary outcome (all-cause death, nonfatal myocardial infarction, and nonfatal stroke) in individuals in the CCB strategy treatment group (HR (95% CI) = 8.58 (2.06–35.8) (Figure 32.7), but not in the β-blocker strategy treatment group (HR (95% CI) = 2.31 (0.82–6.55) (Figure 32.7) [153]. This suggests that treatment with the β-blocker strategy had a protective role in individuals with the Ser49-Arg389 haplotype.

The interaction of genetic variants in *ADRB1* and β-blocker treatment was also investigated in a population-based case-control study of pharmacologically treated hypertensive patients [154]. Cases were identified with incident myocardial infarction or ischemic stroke and were matched to controls; genetic variation throughout *ADRB1* was examined for association with risk of adverse cardiovascular outcomes and interaction between outcomes and hypertensive treatment. Two SNPs (rs17875422 and rs2429511), located in the promoter region of *ADRB1*, showed a significant

TABLE 32.3 Clinical Trials in Treatment-Related Antihypertensive Outcome Pharmacogenomic Studies

Full Name	Acronym	Genetic Samples	Randomized Drug Classes	Average Follow-Up Time (Years)
International Verapamil-SR Trandolapril Study	INVEST [73]	5979	BB, CCB	2.7
Nordic Diltiazem Study	NORDIL [74]	4196	BB, CCB	4.5
Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial	ALLHAT [72]	39,114	CCB, ACEI, TD, A1B	4.9[a]
Losartan Intervention for Endpoint Reduction in Hypertension Study	LIFE [152]	3503	BB, ARB	4.8

[a] *4.9 years was the average follow-up in the CCB, ACEI, and TD arms; the A1B arm was stopped early with an average follow-up of 3.3 years.*

BB—β-blocker; CCB—calcium channel blocker; ACEI—ACE inhibitor; TD—thiazide diuretic; A1B—α-1-blocker; ARB—angiotensin receptor blocker.

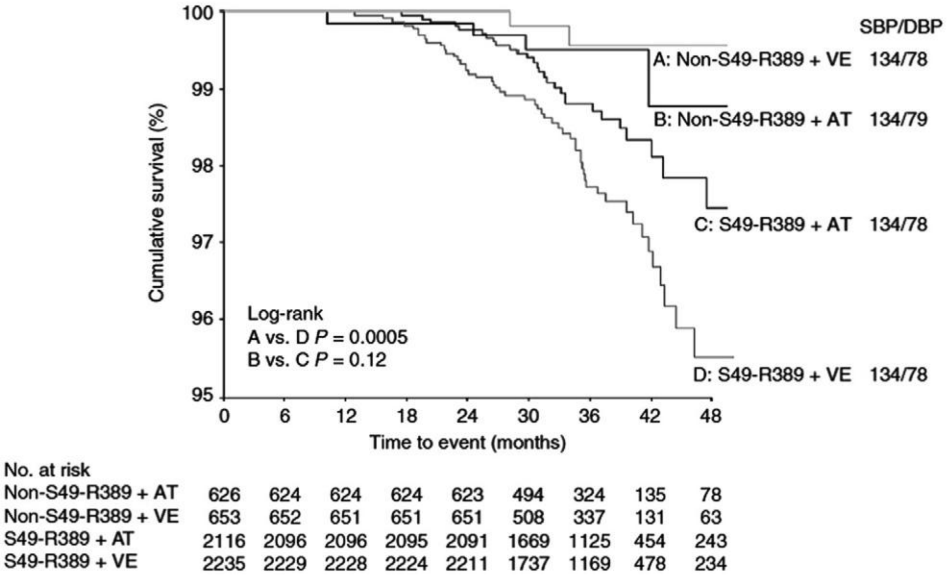


FIGURE 32.7 All-cause mortality and mean on-treatment blood pressure by ADRB1 Ser49-Arg389 haplotype and atenolol/verapamil-SR therapy. HR—hazard ratio; CI—confidence interval; S49-R389, Ser49-Arg389 haplotype; AT—atenolol; VE—verapamil-SR; SBP—systolic blood pressure; DBP—diastolic blood pressure. *Source: From Pacanowski et al.* [153] *with permission.*

interaction between adverse cardiovascular outcomes and β-blocker use [154]. For both SNPs, variants carriers had an increased risk of adverse cardiovascular outcomes when treated with a β-blocker; individuals who were homozygous for the common allele had a decreased risk of adverse cardiovascular outcomes when treated with a β-blocker [154]. Collectively, these two studies, and the functional evidence for the Ser49Gly and Arg389Gly variants, support the role of *ADRB1* in β-blocker treatment-related outcomes.

32.6.2.2 NEDD4L

Another gene that shows evidence of playing a role in both BP response and cardiovascular outcomes is *NEDD4L*. Prior studies in humans have shown that the GG genotype at *NEDD4L* rs4149601 is associated with higher salt sensitivities, lower plasma renin levels, and higher DBP [106–108]. These data suggest individuals with the G allele at rs4149601 may respond better to thiazide diuretics, and have better outcomes.

In NORDIL, this suggestion was confirmed. Those who carried the G allele at rs4149601 and were treated with the β-blocker/diuretic strategy had better cardiovascular outcomes compared to those with the AA genotype (OR (95% CI) = 0.52 (0.36–0.74), $p < 0.0001$) (Figure 32.8) [109]. However, for those individuals treated with diltiazem, there were no significant differences in cardiovascular outcome resulting from the rs4149601 genotype ($p = 0.77$) (Figure 32.8) [109].

In INVEST, rs4149601 was also examined in relation to adverse cardiovascular outcomes and diuretic use (HCTZ). Among individuals treated with HCTZ, there was no evidence of a difference in adverse cardiovascular outcomes by genotype in whites ($p = 0.477$) (Figure 32.8) [110]. However, in whites not treated with HCTZ, there was a significant increase in adverse cardiovascular events in those with the AG genotype as compared to the AA genotype (OR (95% CI) = 10.65; 1.18–96.25); $p = 0.022$ (Figure 32.8), and this trend remained when the AG and GG genotypes were combined (OR (95% CI) = 8.94 (0.99–80.54) $p = 0.051$ (Figure 32.8) [110]. However, there was no replication in Hispanics or African-Americans in INVEST [110]. These two pharmacogenomic studies, the prior associations of *NEDD4L* with hypertension, and the functional implications of rs4149601 all contribute to *NEDD4L*'s importance in thiazide diuretic treatment-related outcomes, with the data suggesting that individuals who carry the G allele may have more favorable outcomes with thiazide diuretic treatment.

32.6.2.3 CACNB2

CACNB2, the β_2 subunit of the voltage-gated L-type calcium channel, is responsible for targeting cell-surface expression of the α_1 pore-forming subunit, the binding site for all currently available CCBs. *CACNB2* has also been associated in GWASs with BP and hypertension [47,155]. It is the only gene to date with two independent signals observed: one in the 3′ end of the gene and one in the 5′ end [47,155].

CACNB2 was also studied in relation to cardiovascular outcomes in INVEST. A SNP that lies in the promoter region of *CACNB2*, rs2357928, was found to have a significant treatment interaction in whites (interaction $p = 0.002$)

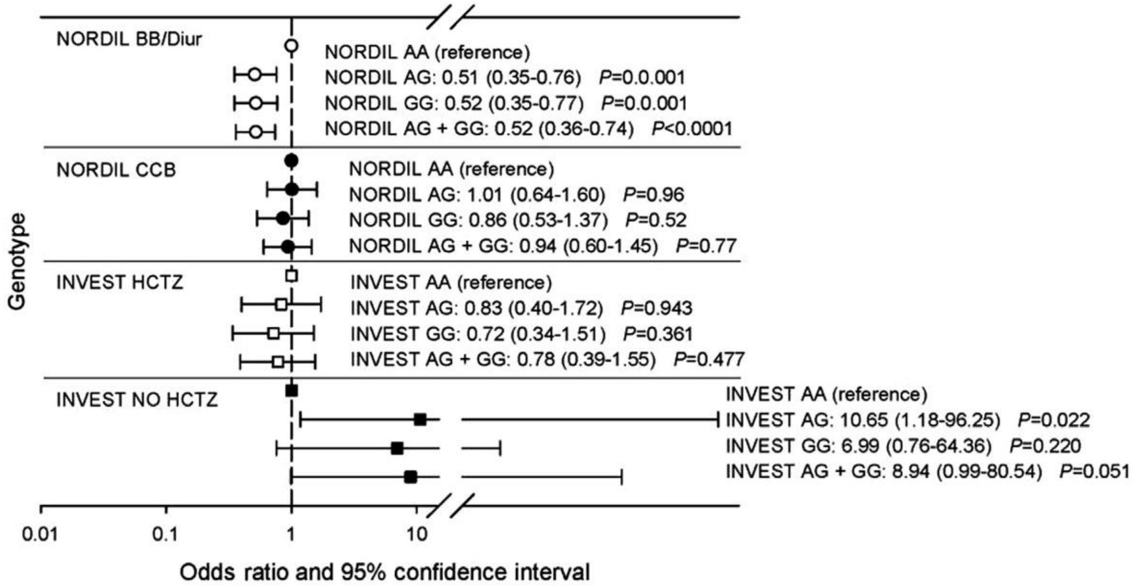


FIGURE 32.8 **Comparison of genotypes.** Impact of *NEDD4L* rs4149601 genotypes AG and GG compared to AA (reference): in patients treated with β-blockers/diuretics (BB/Diur) or diltiazem (CCB) on the risk of receiving a primary end point in NORDIL; and in patients treated with hydrochlorothiazide (HCTZ) or without hydrochlorothiazide (NO HCTZ) on the risk of receiving the primary outcome in INVEST. *Source: Data from Svensson-Farbom et al.* [109] *and McDonough et al.* [110].

(Figure 32.9) [125]. This interaction remained significant after false discovery rate (FDR) adjustment (FDR-adjusted interaction p = 0.024). In whites, those with the GG genotype were found to be at higher risk for the primary outcome when treated with a CCB strategy compared to a β-blocker strategy (HR (95% CI) = 2.35 (1.19–4.66), p = 0.0143 (Figure 32.9) [125].

In contrast, in whites who were A carriers, there was no evidence of a difference in outcomes by treatment strategy (Figure 32.9). This association did not appear to be mediated through differences in BP response. Additionally, there was a trend observed in the same direction for treatment interaction in Hispanics at rs2357928 (interaction p = 0.069) [125]. It was also found that the G-allele at rs2357928 significantly increases promoter activity, as indicated by a luciferase assay [125]. Together these data suggest that, in addition to a likely role in BP and hypertension, variants in *CACNB2* may also influence treatment-related outcomes in patients with hypertension.

32.6.2.4 *CACNA1C*

CACNA1C encodes the α1c subunit of the L-type calcium channel. This subunit is the major pore-forming subunit of the L-type calcium channel and is the binding site for all currently available CCBs. Variants in *CACNA1C* were associated with SBP, DBP, and hypertension in the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium when examining antihypertensive drug targets [121]. Additionally, in a whole-exome sequencing and bioinformatics analysis for autosomal dominant long QT syndrome, a novel variant in *CACNA1C* (Pro857Arg) was identified, and was shown to cause a gain-of-function mutation leading to increased expression of the channel [156]. This gene was evaluated for pharmacogenomic effect in INVEST. A synonymous SNP at codon 1835, rs1051375, was found to have a significant interaction with treatment strategy in the combined cohort containing whites, Hispanics, and African-Americans (interaction p = 0.0001) [124]. Individuals with the AA genotype had lower rates of the primary outcome (all-cause death, nonfatal myocardial infarction, and nonfatal stroke) in the CCB strategy (Figure 32.10) [124], whereas individuals with the GG genotype had lower rates of the primary outcome in the β-blocker strategy (Figure 32.10) [124]. This treatment interaction remained when looking at whites alone (interaction p = 0.002) [124]. These pharmacogenomic results and the associations of SNPs in *CACNA1C* with BP and other cardiovascular traits support *CACNA1C*'s role in treatment-related cardiovascular outcomes.

32.6.2.5 *Fibrinogen-beta*

The fibrinogen-beta (*FGB*) gene encodes the fibrinogen-β chain, one of the three polypeptide chains that make up the coagulation factor fibrinogen. Higher plasma fibrinogen levels have been associated with coronary heart disease, stroke, and mortality [157]. Additionally, the A allele at the *FGB*-455 G>A variant has been associated with higher levels of fibrinogen [158]. It is also known that ACE inhibitors

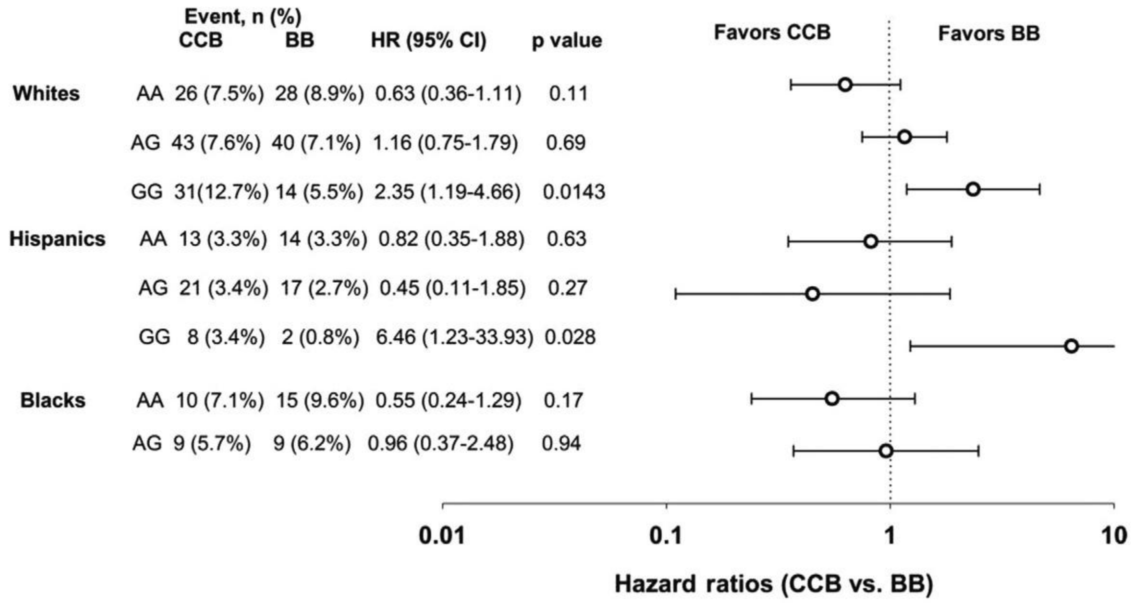


FIGURE 32.9 **Pharmacogenetic association of *CACNB2* rs2357928 and treatment strategies in the three race/ethnicity groups in INVEST.** Hazard ratios are shown in log scale. CCB—calcium channel blocker treatment strategy; BB—β-blocker treatment strategy. *Source: From Niu et al.* [125] *with permission.*

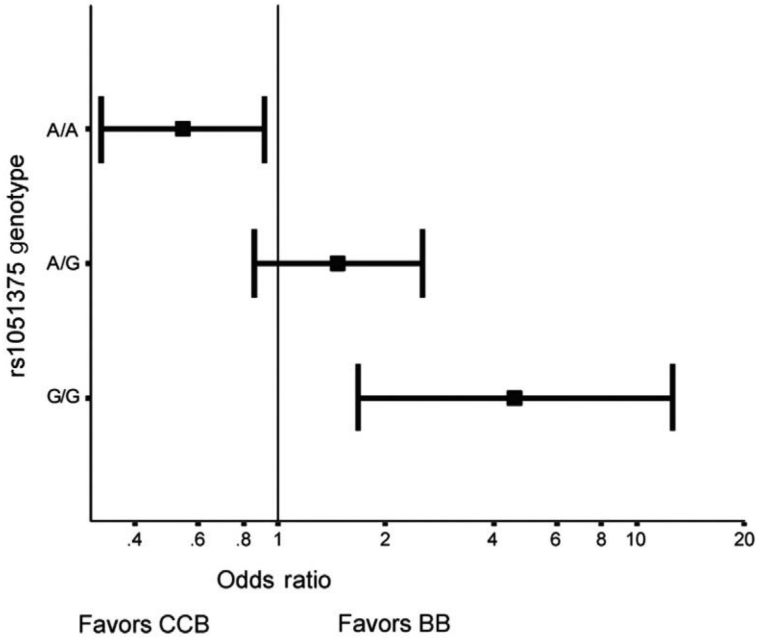


FIGURE 32.10 **Adjusted ORs and 95% CIs for treatment strategy by the *CACNA1C* rs1051375 genotype in INVEST.** CCB—calcium channel blocker treatment strategy; BB—β-blocker treatment strategy. *Source: From Beitelshees et al.* [124] *with permission.*

lower plasma fibrinogen concentrations [159]. The pharmacogenomic relationship between the *FGB*-455 G>A variant and ACE inhibitors was investigated with cardiovascular outcomes (stroke, heart failure, all-cause mortality, and end-stage renal disease) in the GenHAT (Genetics of Hypertension Associated Treatment) study, the genetic substudy of ALLHAT [160]. It was found that GG homozygotes had a higher risk of stroke and mortality when treated with lisinopril compared to amlodipine (stroke: HR = 1.38, $p < 0.001$; mortality: HR = 1.12, $p = 0.02$), and that A allele carriers had a slightly lower risk (interaction $p = 0.03$ and 0.04 for stroke and mortality, respectively) [160]. It was also found that GG homozygotes had a higher risk of mortality and end-stage renal disease when treated with lisinopril compared to chlorthalidone (mortality: $HR = 1.05$, $p = 0.23$; end-stage renal disease: $HR = 1.27$, $p = 0.08$), and that A allele carriers had a lower risk (interaction $p = 0.03$ for both mortality and end-stage renal disease) [160]. These results, coupled with the functional evidence of association between fibrinogen levels and ACE inhibitors, and fibrinogen levels and the *FGB*-455 G>A variant, suggest that *FGB* may be involved in ACE inhibitor treatment-related outcomes.

32.6.2.6 Genetic Risk Scores: SIGLEC12, A1BG, and the Selectin Region

One of the issues in hypertension pharmacogenomics is how to put the results together into something that is clinically meaningful, as there are many single-gene associations, each contributing a smaller effect. One method that can be used to combine the results of many associations is the building of a risk score, where each risk allele or genotype is assigned a point value and the sum of the point values across all associated SNPs becomes the test variable. This approach was used in an analysis of nonsynonymous SNPs on the Illumina HumanCVD chip in INVEST [161]. First, single SNPs were identified that were associated with a treatment interaction when looking at adverse cardiovascular outcomes [161]. Two SNPs were found that were associated with treatment-related cardiovascular outcomes in INVEST whites and Hispanics: *SIGLEC12* rs16982743 (Gln29Stop) and *A1BG* rs893184 (His52Arg) (combined interaction $p = 0.0038$ and 0.0036, respectively) [161]. Additionally a region on chromosome 1 was identified that showed associations with treatment interaction in both INVEST whites and Hispanics. This region encompassed *F5*, *SELP*, and *SELE*; through additional analyses, rs4525 (His865Arg) in *F5* was selected as the best signal to represent the region.

A genetic risk score constructed from rs16982743, rs893184, and rs4525 was significantly associated with treatment-related adverse cardiovascular outcomes in whites and Hispanics from INVEST and in NORDIL (meta-analysis interaction $p = 2.39 \times 10^{-5}$) (Figure 32.11) [161]. In individuals with a genetic risk score of 0 or 1, CCB treatment was associated with lower risk of adverse cardiovascular outcomes compared to β-blocker treatment (OR (95% CI) = 0.60 (0.42–0.86) (Figure 32.11). In those with a genetic risk score of 2–3, CCB treatment was associated with higher risk of adverse cardiovascular outcomes compared to β-blocker treatment, OR (95% CI) = 1.31 (1.08–1.59) (Figure 32.11) [161].

These gene regions all are related to immune cell trafficking and/or are in the immunoglobulin superfamily. *SIGLEC12* encodes the sialic acid–binding Ig-like lectin 12, a single-pass, type I transmembrane protein belonging to

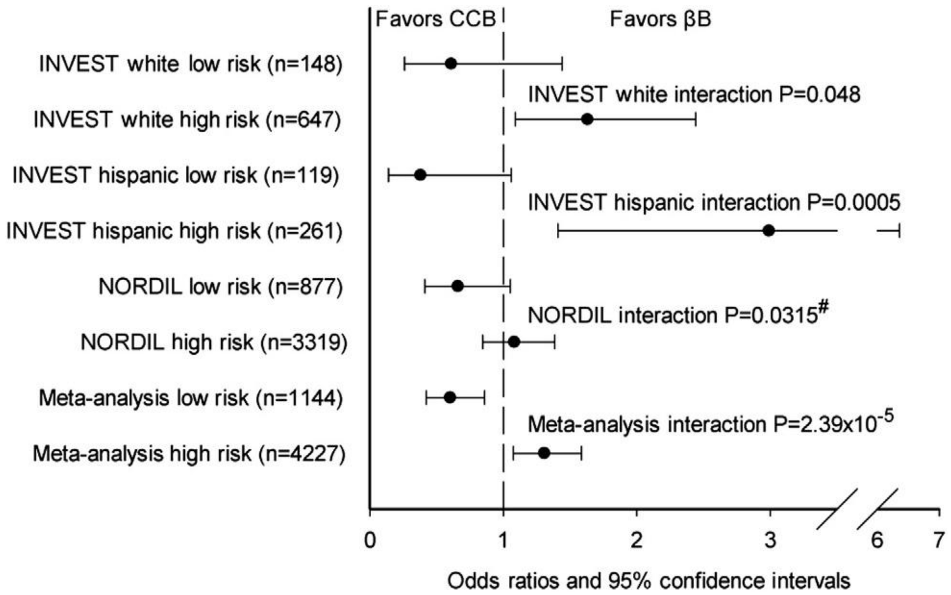


FIGURE 32.11 **Risk-score pharmacogenomic association with the primary outcome in INVEST.** Risk score calculated from *SIGLEC12* rs16982743, *A1BG* rs893184, and *F5* rs4525. One point was given for each genotype that conferred higher risk in the CCB arm/group versus the βB arm/group: rs16982743 GG (Gln/Gln), rs893184 GG (Arg/Arg), and rs4525 AA (His/His). Low Risk defined as a risk score = 0 or 1; High Risk defined as a risk score = 2 or 3. #: One-sided *P*-value based on one-sided hypothesis in replication cohort. CCB: Calcium channel blocker treatment strategy. βB: β-blocker treatment strategy. *Source: Used with permission from McDonough et al.* [161].

the immunoglobulin superfamily. It is expressed on the epithelial cell surface and on some macrophages. It has been suggested that *SIGLEC12* could be involved in the negative regulation of macrophage signaling by functioning as an inhibitory receptor [162]. Additionally, other members of the *SIGLEC* family have been shown to be expressed abundantly on macrophages recruited during the pathogenesis of atherosclerosis [163]. The alpha 1B-glycoprotein precursor (*A1BG*) encodes a plasma glycoprotein with unknown function. The A1BG protein is a member of the immunoglobulin superfamily, and has been shown to form a complex with CRISP-3 (cysteine-rich secretory protein 3), which is present in neutrophilic granulocytes and is thought to play a role in innate immunity [164].

The last region identified from the treatment interaction analyses contains signals in three genes: *F5*, coagulation factor V; *SELP*, selectin P; and *SELE*, selectin E. *F5* encodes an essential cofactor in the blood coagulation cascade, participating as half of the complex that activates prothrombin to thrombin during the blood-clotting process [165]. *SELP* and *SELE* are members of the selectin family of cell adhesion molecules, and both mediate leukocyte rolling, a process essential for the initiation of atherosclerosis [166]. P-selectin is in stored platelets and endothelial cells and, upon activation, redistributes to the plasma membrane; E-selectin, on the other hand, is only expressed on endothelial cells [166]. Many prior genetic associations have been seen with *F5*, *SELP*, *SELE*, and CVD [167–169].

These results suggest cardiovascular outcomes may differ based on the *SIGLEC12*, *A1BG*, and *F5* genotypes and antihypertensive treatment strategy. Such specific genetic associations and the use of a genetic risk score provide insight into a potential approach to selecting personalized antihypertensive treatment.

32.6.2.7 *ADD1*

There are genes that initially showed evidence of a positive pharmacogenomic association, but for which subsequent studies have been unable to replicate results. This is the case with *ADD1*, the α-adducin gene. *ADD1* is a ubiquitously expressed protein that is part of the cytoskeleton and may modulate ion transport. One SNP in *ADD1*, Gly460Trp (rs4961), which was previously associated with hypertension and salt sensitivity [95], has been studied many times for a pharmacogenomic effect. A population-based case-control study found that, in Trp460 variant carriers, diuretic use was associated with a significant decrease in the risk of myocardial infarction or stroke compared to all other antihypertensive therapies (OR (95% CI) = 0.49 (0.32–0.77)) [170]. However, this was also studied in GenHAT [171], INVEST [172], PHARMO [173], and LIFE [139], and was not replicated. Taken together, these results suggest that there is no pharmacogenomic effect between *ADD1* Gly460Trp, cardiovascular outcomes, and diuretic use.

32.6.2.8 *ACE*

ACE is another gene with strong evidence for no pharmacogenomic effect. The *ACE* I/D variant, a 287-bp insertion/deletion in intron 16, has been extensively studied: individuals with the D allele have higher plasma ACE activity [132]. The interaction between the *ACE* I/D genotype, antihypertensive treatment, and risk of adverse cardiovascular outcomes was studied in GenHAT [174]. There was no difference by *ACE* I/D genotype, antihypertensive treatment (diuretic, CCB, ACE inhibitor, or α-blocker), and cardiovascular outcomes ($p \geq 0.26$) [174].

Two other studies saw similar results, observing no evidence of a pharmacogenomic effect for adverse cardiovascular outcomes by *ACE* I/D genotype [139,175]. In the LIFE study, there were no differences between *ACE* I/D genotype, cardiovascular events, and treatment with an ACE inhibitor or β-blocker [139]. In a hypertension case-control cohort from the population-based Rotterdam study, there were also no differences observed between *ACE* I/D genotype, risk of myocardial infarction or stroke, and ACE inhibitor or β-blocker use [175]. However, in a population-based registry of pharmacy records linked to hospital discharge records (PHARMO), significant treatment interaction was observed between *ACE* rs4341 (which is in linkage disequilibrium with *ACE* I/D), ACE inhibitor treatment, and the risk of myocardial infarction [176]. Individuals who were ACE inhibitor users with the rs4341 G allele (which is inherited with the *ACE* D allele) had a greater reduction in risk for myocardial infarction compared to those with the CC genotype (synergy index for interaction (SI) = 0.58, 95% CI, 0.35–0.95) [176]. While there were significant results in one study, the other three studies did not see an association between antihypertensive treatment outcomes and *ACE* I/D genotype, which suggests that there is no pharmacogenomic effect.

32.6.3 Additional Genes with Evidence of Association with Treatment-Related Outcomes

Other genes evincing association with treatment-related outcomes do not have replication but are reasonable candidate genes, many antihypertensive drug targets, or hypertension genes. In many cases, these have been evaluated in only one study, but their biological plausibility makes them worth mention. These include *KCNMB1*, *NR1H3*, *MMP9*, *MMP12*, *NOS3*, *NPPA*, *RYR3*, *AGT*, and *AGTR1*. Many of these were combined through the use of a gene panel in GenHAT [177].

32.6.3.1 *KCNMB1*

The interaction between *KCNMB1* nonsynonymous SNPs, Glu65Lys and Val110Leu, and treatment strategy for long-term cardiovascular outcome (all-cause death, nonfatal myocardial infarction, and nonfatal stroke) was studied in INVEST [131]. The results showed that at the Val110Leu variant, Leu110 variant carriers had a 32% reduced risk of the primary outcome compared to noncarriers when both treatment groups were combined [131]. While the treatment interaction term was not significant with the Val110Leu variant, the association was limited to those individuals randomized to the CCB strategy (HR (95% CI) = 0.55 (0.31–0.97)), and was not observed in those randomized to the β-blocker strategy (HR (95% CI) = 0.84 (0.51–1.41)) [131].

32.6.3.2 Liver X Receptor Alpha

NR1H3 encodes for liver X receptor-α (LXRA), a nuclear receptor that has been associated with body mass index and cholesterol levels. Evidence has also shown that ligands of liver X receptors decrease RAAS activation [178]. Taken together, this suggests that variants in *NR1H3* may contribute to adverse cardiovascular outcomes. Seven tag SNPs in *NR1H3* were evaluated for association with adverse cardiovascular events in INVEST [179]. A significant treatment interaction was observed with rs2279238, where T carriers had increased risk of adverse cardiovascular outcomes in the CCB strategy compared to the β-blocker strategy (OR (95% CI) = 2.86 (1.50–5.46), Interaction p = 0.0015) [179].

32.6.3.3 MMP9 and MMP12

MMP9 and *MMP12* are two matrix metalloproteinases. MMP-9, gelatinase-B, functions in the degradation of type IV and V collagens; elevated levels of MMP-9 have been observed in individuals with hypertension. MMP-12 is also known as the macrophage metalloelastase and is up regulated in atherosclerotic lesions. Two SNPs in *MMP9* (Arg279Gln, rs17576; Arg668Gln, rs2274756) and two SNPs in *MMP12* (Asn122Ser, rs652438; −82A>G, rs2276109) were investigated in the GenHAT study to determine whether hypertensive participants randomized to a diuretic (chlorthalidone), a CCB (amlodipine), or an ACE inhibitor (lisinopril) had different cardiovascular outcome rates.

Pharmacogenomic associations were seen for Arg279Gln, Arg668Gln, and Asn122Ser with numerous cardiovascular outcomes. In general, at *MMP9* Arg279Gln, modest pharmacogenomic effects were seen for the composite cardiovascular disease outcome and for combined coronary heart disease, with individuals with the GG genotype having lower risk when randomized to the diuretic versus the CCB or the ACE inhibitor. This is in comparison to individuals who were A carriers [180]. At *MMP9* Arg668Gln, lower HRs were observed for individuals with the AA genotype for most diuretic treatment outcomes when compared to CCB treatment. At the *MMP12* Asn122Ser SNP, higher HRs were seen in individuals with at least one G allele who were on the ACE inhibitor compared to the CCB for both coronary heart disease and the composite cardiovascular disease outcome. However, there were no pharmacogenomic effects observed with the −82A>G variant [180].

32.6.3.4 NOS3

NOS3, nitric oxide synthase 3, is the endothelial nitric oxide synthase that catalyzes nitric oxide production. Nitric oxide is a signaling molecule involved in vasodilation and vascular smooth-muscle relaxation. Decreases in nitric oxide lead to hypertension and atherosclerosis [181], whereas overproduction leads to cell damage. Variants in *NOS3* have been associated with various cardiovascular diseases, including myocardial infarction [182] and coronary artery disease [183]. Additionally, CCB's mechanism of inhibiting calcium influx into smooth-muscle cells leads to vasorelaxation and may act through activation of *NOS3* [184].

The theory that variants in *NOS3* may be associated with different cardiovascular outcomes in individuals on a CCB treatment strategy was tested in the GenHAT study [185]. Three SNPs in *NOS3*—rs3918226, rs1800779, and rs1799983—were tested for a pharmacogenetic effect on cardiovascular outcomes. Significant findings were observed with stroke and all-cause mortality [185]. At rs3918226 (*NOS3*-690 C>T), T allele carriers had a lower risk for stroke when treated with a CCB compared to an ACE inhibitor (HR (95% CI) = 0.49 (0.31–0.80), interaction p = 0.04) [185]. Also, at rs1799983 (*NOS3* Glu298Asp), Asp carriers (T carriers) had a lower risk of all-cause mortality when treated with a CCB compared to an ACE inhibitor (HR (95% CI) = 0.85 (0.75–0.97), interaction p = 0.04) [185].

32.6.3.5 NPPA

The atrial natriuretic precursor A gene, *NPPA*, encodes the precursor for atrial natriuretic polypeptide, which acts as a diuretic by controlling extracellular fluid volume and electrolyte homeostasis [186]. Prior associations have been seen between SNPs in *NPPA* and adverse cardiovascular outcomes such as stroke [187] and coronary artery disease [188]. Two of these SNPs, *NPPA* G664A and *NPPA* T2238C, were studied in GenHAT to determine if individuals treated with a diuretic would have different outcomes based on their genotype when compared to individuals treated with a CCB or an ACE inhibitor [189]. A pharmacogenetic effect was observed at *NPPA* T2238C—C allele carriers had lower rates of coronary heart disease and stroke when treated with a diuretic compared to a CCB; individuals with the TT genotype had lower rates when treated with a CCB compared to a diuretic (interaction p = 0.03 and 0.01, respectively, for coronary heart disease and stroke) [189].

32.6.3.6 RYR3

Genes involved in the maintenance of calcium homeostasis have also been considered as candidate genes for hypertension pharmacogenomics. One is the ryanodine receptor 3 gene (*RYR3*), an intracellular calcium channel expressed in human arterial endothelial cells which mediates the mobilization of stored calcium to initiate muscle contraction in cardiac and skeletal muscle [190]. SNPs in *RYR3* have been associated with carotid intima-media thickness [191]. Additionally, evidence associates both *RYR3* and CCBs with heart failure [192,193]. Using a case-only design, GenHAT tested three SNPs in *RYR3* for a pharmacogenomic effect between antihypertensive treatment and adverse cardiovascular outcomes [194]. There was evidence of a pharmacogenomic effect with two of the SNPs, rs877087 and rs2077268, and heart failure. At rs877087, TT homozygotes fared worse when treated with a CCB compared to a diuretic (TT cases = 26% for CCB and 19% for diuretic); also, CC homozygotes fared better when treated with a CCB compared to a diuretic (CC cases = 32% for CCB and 29% for diuretic, interaction $p = 0.005$) [194]. At rs2077268, similar trends were seen: minor allele homozygotes, TT, did worse on a CCB compared to a diuretic (TT cases = 5% for CCB and 2% for diuretic), and major allele homozygotes, CC, did better on a CCB compared to a diuretic (CC cases = 69% for CCB and 72% for diuretic, interaction $p = 0.01$) [194].

32.6.3.7 AGT

The pharmacogenomic effect of the *AGT* Met235Thr variant was investigated in the Rotterdam study and in PHARMO [176,195]. In the Rotterdam population-based case-control cohort, ACE inhibitor users with one copy of the Thr allele (Thr/Thr or Met/Thr genotype) had an increased risk of myocardial infarction compared to ACE inhibitor users with the Met/Met genotype (SI (95% CI) = 4.00 (1.32–12.11)) [195]. However, the PHARMO study found no association between the Met235Thr genotype, ACE inhibitor use, and myocardial infarction risk [176].

32.6.3.8 AGTR1

SNPs in *AGTR1* were investigated in both the Rotterdam study and PHARMO for any association with antihypertensive treatment and adverse cardiovascular outcomes [175,176]. The *AGTR1* 573 C/T polymorphism (rs5182) was investigated in a population based case-control cohort from the Rotterdam study. It was found that individuals currently using ACE inhibitors with the CT or CC genotype at rs5182 had a reduced risk of myocardial infarction as compared to those currently using ACE inhibitors with the TT genotype [175]. There was no significant interaction observed with the risk of stroke [175].

The PHARMO study looked at the pharmacogenomic effect of the *AGTR1* 1166 A/C polymorphism in relation to antihypertensive treatment and myocardial infarction risk. In ACE inhibitor users who were prescribed low daily doses, the risk of myocardial infarction was reduced in A carriers compared to those with the CC genotype (SI (95% CI) = 3.67 (1.18–11.4)) [176].

32.6.3.9 Gene Panels

Another method of combining multiple associations is the gene panel, which stratifies patients into different treatment groups. This approach was used in GenHAT to identify subgroups of individuals at high and low risk of coronary heart disease based on antihypertensive treatment [196]. The initial panel contained 78 SNPs in hypertension pharmacogenomic candidate genes; a specific panel was identified to stratify individuals into each of the four treatment groups: diuretic, CCB, ACE inhibitor, and α-blocker [196]. The panels in the specific treatment groups were then used to see if including the SNP data improved the prediction of coronary heart disease.

Overall, the panels identified between the four treatment groups had little overlap, but they did improve the prediction of coronary heart disease in each treatment group at a statistically significant level [196]. For the diuretic, the SNPs included were *NOS3* rs3918226, *SELE* rs5361, *ICAM1* rs1799969, *AGT* rs5051, and *GNAS* rs7121. For the CCB, the following SNPs were identified: *MMP1* rs1799750, *F5* rs6025, *NPPA* rs5065, *PDE4D* rs6450512, and *MMP9* rs2274756. The SNPs on the panel for the ACE inhibitor were *AGT* rs5051, *PON1* rs705379, *MMP12* rs652428, *F12* rs1801020, *GP1BA* rs6065, and *PDE4D* rs27653. The SNPs included for the α-blocker panel were *F2* rs1799963, *PAI1* rs1799768, *MMP7* rs11568818, *AGT* rs5051, *ACE* rs4343, and *MMP2* rs243865 [196]. Each of the four panels was also tested for a pharmacogenomic effect; evidence was seen in three of them (diuretic, CCB, and α-blocker) [196].

Overall, the top signals so far identified as having an association with treatment interaction are in strong biologically plausible genes; they have evidence of replication and/or associations with related phenotypes. Additionally, many of the SNPs have been shown to be functional. These top genes include *ADRB1*, *NEDD4L*, *CACNB2*, *CACNA1C*, and *FGB* (identified in candidate gene approaches) and *SIGLEC12*, *A1BG*, and *F5* (identified by genome spanning). While all of these results are convincing, additional replication is necessary and studies need to be conducted to determine how these results can be used in the clinical setting to stratify patients to a specific antihypertensive treatment regimen or strategy based on genetic information.

32.7 FUTURE PERSPECTIVES

Although presently there are many genes with compelling data for both BP response and treatment-related outcomes, more work needs to be done. Many of the genes highlighted in this chapter need additional replication and functional studies, especially those identified through GWASs, to further elucidate their role in hypertension pharmacogenomics and their mechanisms of action. Additionally, studies must be carried out to determine how these data can be incorporated into patient care.

The first step in integrating results in the clinical setting will be determining which SNPs/genes have enough evidence to be clinically actionable. Once there is enough evidence, methods of combining results will have to be investigated. Two approaches have so far been used to combine multiple single SNP/gene findings to make them more useful in a clinical setting. The first, genetic risk scores, has been used in BP response [197] and treatment-related outcomes [161]. Genetic risk scores are more commonly seen in disease genetics [198,199], but their use in pharmacogenomics is valuable for future clinical translation. This is especially true for hypertension pharmacogenomics, as multiple genetic signals are likely involved. Risk scores are typically constructed from prior association signals and tested in an independent population [198,199]. In both cases of the use of genetic risk scores in hypertension pharmacogenomics, the scores were constructed in the discovery population but were confirmed through permutation testing [197] or through replication in an independent population [161]. The second method of combining multiple SNPs/genes employs a gene panel to stratified subjects into a treatment category. This method was used in GenHAT to identify risk groups for coronary heart disease [177]. Both methods rely on combining multiple signals into one clinically useful variable.

In the future, it is likely that whole-genome SNP or sequence data will be available in a patient's medical record [200]. When this occurs, the implementation of genetic data in patient care will not depend on a genetic test result and there will be no need to conduct multiple tests to determine what drug the patient responds to. At present, a single SNP chip would be most useful as a means of stratifying patients to the best antihypertensive treatment option. This way only a single test would be ordered. However, difficulties remain, as we have many new discoveries every year, and the level of evidence differs according to antihypertensive drug class.

There are still many unanswered questions in hypertension pharmacogenomics, and many regions of the genome have yet to be interrogated. There are many ongoing GWASs for both BP response and treatment-related outcomes. For BP response, these include PEAR [75], GENRES [77], GERA [76,86], and NORDIL [74], as well as UNISA and INPH (Italian Network for the Pharmacogenomics of Hypertension). For treatment-related outcomes, there are INVEST [73], NORDIL [74], ASCOT [71], and SPS3 [201]. There is also ALLHAT's GenHAT substudy [72], which, although not doing a GWAS at this time, may be useful for replication. The completion and publication of these studies will further define the potential for clinical implementation in hypertension pharmacogenomics.

The International Consortium for Antihypertensives Pharmacogenomics Studies, or ICAPs (www.pharmgkb.org/page/icaps) was formed in mid-2012 with the goal of facilitating collaboration among research groups and performing meta-analyses for both BP response and treatment-related outcomes. The international collaboration that ICAPs fosters will advance the field of hypertension pharmacogenomics; moreover, the results of the meta-analyses should provide a list of SNPs and genes that have strong evidence for clinical implementation.

All in all, the next decade should see many exciting results in the field of hypertension pharmacogenomics.

REFERENCES

[1] Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Collaboration PS. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet 2002;360(9349):1903–13.

[2] Evans JG, Rose G. Hypertension. Br Med Bull 1971;27(1):37–42.

[3] Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. Lancet 2005;365(9455):217–23.

[4] Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the global burden of disease study 2010. Lancet 2012;380(9859):2224–60.

[5] Leitschuh M, Cupples LA, Kannel W, Gagnon D, Chobanian A. High-normal blood pressure progression to hypertension in the Framingham heart study. Hypertension 1991;17(1):22–7.

[6] Franklin SS, Gustin W, Wong ND, Larson MG, Weber MA, Kannel WB, et al. Hemodynamic patterns of age-related changes in blood pressure. The Framingham heart study. Circulation 1997;96(1):308–15.

[7] Stamler R, Stamler J, Riedlinger WF, Algera G, Roberts RH. Weight and blood pressure. Findings in hypertension screening of 1 million Americans. JAMA 1978;240(15):1607–10.

[8] Frisoli TM, Schmieder RE, Grodzicki T, Messerli FH. Salt and hypertension: is salt dietary reduction worth the effort? Am J Med 2012;125(5):433–9.

[9] He FJ, MacGregor GA. Effect of modest salt reduction on blood pressure: a meta-analysis of randomized trials. Implications for public health. J Hum Hypertens 2002;16(11):761–70.

[10] Kaminer B, Lutz WP. Blood pressure in Bushmen of the Kalahari desert. Circulation 1960;22:289–95.

[11] Truswell AS, Kennelly BM, Hansen JD, Lee RB. Blood pressures of Kung Bushmen in Northern Botswana. Am Heart J 1972;84(1):5–12.

[12] Poulter NR, Khaw KT, Hopwood BE, Mugambi M, Peart WS, Rose G, et al. The Kenyan Luo migration study: observations on the initiation of a rise in blood pressure. BMJ 1990;300(6730):967–72.

[13] Crews DE, Mancilha-Carvalho JJ. Correlates of blood pressure in Yanomami Indians of Northwestern Brazil. Ethn Dis 1993;3(4):362–71.

[14] Carvalho JJ, Baruzzi RG, Howard PF, et al. Blood pressure in four remote populations in the INTERSALT study. Hypertension 1989;14(3):238–46.

[15] Laville M, Lengani A, Sermé D, Fauvel JP, Ouandaogo BJ, Zech P. Epidemiological profile of hypertensive disease and renal risk factors in black Africa. J Hypertens 1994;12(7):839–43. [discussion 45].

[16] Cordain L, Eaton SB, Sebastian A, Mann N, Lindeberg S, Watkins BA, et al. Origins and evolution of the Western diet: health implications for the 21st century. Am J Clin Nutr 2005;81(2):341–54.

[17] Neel JV. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet 1962;14:353–62.

[18] Burt VL, Whelton P, Roccella EJ, Brown C, Cutler JA, Higgins M, et al. Prevalence of hypertension in the US adult population. Results from the third national health and nutrition examination survey, 1988–1991. Hypertension 1995;25(3):305–13.

[19] Nakajima T, Wooding S, Sakagami T, Emi M, Tokunaga K, Tamiya G, et al. Natural selection and population history in the human angiotensinogen gene (AGT): 736 complete AGT sequences in chromosomes from around the world. Am J Hum Genet 2004;74(5):898–916.

[20] Weder AB. Evolution and hypertension. Hypertension 2007;49(2): 260–5.

[21] Young JH, Chang YP, Kim JD, Chretien JP, Klag MJ, Levine MA, et al. Differential susceptibility to hypertension is due to selection during the out-of-Africa expansion. PLoS Genet 2005;1(6):e82.

[22] Brunner HR, Gavras H. Is the renin system necessary? Am J Med 1980;69(5):739–45.

[23] Mongeau JG, Biron P, Sing CF. The influence of genetics and household environment upon the variability of normal blood pressure: the montreal adoption survey. Clin Exp Hypertens A 1986;8(4–5):653–60.

[24] Harrap SB, Stebbing M, Hopper JL, Hoang HN, Giles GG. Familial patterns of covariation for cardiovascular risk factors in adults: the Victorian family heart study. Am J Epidemiol 2000;152(8):704–15.

[25] Hunt SC, Williams RR, Barlow GK. A comparison of positive family history definitions for defining risk of future disease. J Chronic Dis 1986;39(10):809–21.

[26] Bochud M, Bovet P, Elston RC, Paccaud F, Falconnet C, Maillard M, et al. High heritability of ambulatory blood pressure in families of East African descent. Hypertension 2005;45(3):445–50.

[27] Havlik RJ, Garrison RJ, Feinleib M, Kannel WB, Castelli WP, McNamara PM. Blood pressure aggregation in families. Am J Epidemiol 1979;110(3):304–12.

[28] Hottenga JJ, Boomsma DI, Kupper N, Posthuma D, Snieder H, Willemsen G, et al. Heritability and stability of resting blood pressure. Twin Res Hum Genet 2005;8(5):499–508.

[29] Kupper N, Willemsen G, Riese H, Posthuma D, Boomsma DI, de Geus EJ. Heritability of daytime ambulatory blood pressure in an extended twin design. Hypertension 2005;45(1):80–5.

[30] Snieder H, Harshfield GA, Treiber FA. Heritability of blood pressure and hemodynamics in African- and European-American youth. Hypertension 2003;41(6):1196–201.

[31] Fava C, Burri P, Almgren P, Groop L, Hulthén UL, Melander O. Heritability of conventional and ambulatory blood pressures. A study in twins. Hypertension 1995;26(6 Pt 1):919–24.

[32] Fava C, Burri P, Almgren P, Groop L, Hulthén UL, Melander O. Heritability of ambulatory and office blood pressure phenotypes in Swedish families. J Hypertens 2004;22(9):1717–21.

[33] Caulfield M, Munroe P, Pembroke J, Samani N, Dominiczak A, Brown M, et al. Genome-wide mapping of human loci for essential hypertension. Lancet 2003;361(9375):2118–23.

[34] Kurtz TW, Morris RC. Dietary chloride as a determinant of "sodium-dependent" hypertension. Science 1983;222(4628):1139–41.

[35] Whitescarver SA, Ott CE, Jackson BA, Guthrie GP, Kotchen TA. Salt-sensitive hypertension: contribution of chloride. Science 1984;223(4643):1430–2.

[36] Lifton RP, Gharavi AG, Geller DS. Molecular mechanisms of human hypertension. Cell 2001;104(4):545–56.

[37] Lifton RP, Dluhy RG, Powers M, Rich GM, Cook S, Ulick S, et al. A chimaeric 11 beta-hydroxylase/aldosterone synthase gene causes glucocorticoid-remediable aldosteronism and human hypertension. Nature 1992;355(6357):262–5.

[38] Lafferty AR, Torpy DJ, Stowasser M, Taymans SE, Lin JP, Huggard P, et al. A novel genetic locus for low renin hypertension: familial hyperaldosteronism type II maps to chromosome 7 (7p22). J Med Genet 2000;37(11):831–5.

[39] Cerame BI, New MI. Hormonal hypertension in children: 11beta-hydroxylase deficiency and apparent mineralocorticoid excess. J Pediatr Endocrinol Metab 2000;13(9):1537–47.

[40] New MI, Levine LS. Hypertension of childhood with suppressed renin. Endocr Rev 1980;1(4):421–30.

[41] Wilson FH, Disse-Nicodème S, Choate KA, Ishikawa K, Nelson-Williams C, Desitter I, et al. Human hypertension caused by mutations in WNK kinases. Science 2001;293(5532):1107–12.

[42] Hansson JH, Nelson-Williams C, Suzuki H, Schild L, Shimkets R, Lu Y, et al. Hypertension caused by a truncated epithelial sodium channel gamma subunit: genetic heterogeneity of Liddle syndrome. Nat Genet 1995;11(1):76–82.

[43] Shimkets RA, Warnock DG, Bositis CM, Nelson-Williams C, Hansson JH, Schambelan M, et al. Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. Cell 1994;79(3):407–14.

[44] Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, et al. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet 2008;40(5):592–9.

[45] John SW, Krege JH, Oliver PM, Hagaman JR, Hodgin JB, Pang SC, et al. Genetic decreases in atrial natriuretic peptide and salt-sensitive hypertension. Science 1995;267(5198):679–81.

[46] Schillinger KJ, Tsai SY, Taffet GE, Reddy AK, Marian AJ, Entman ML, et al. Regulatable atrial natriuretic peptide gene therapy for hypertension. Proc Natl Acad Sci USA 2005;102(39):13789–94.

[47] Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasmnan DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011;478(7367):103–9.

[48] Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet 2009.

[49] Newton-Cheh C, Larson MG, Vasan RS, Levy D, Bloch KD, Surti A, et al. Association of common variants in NPPA and NPPB with circulating natriuretic peptides and blood pressure. Nat Genet 2009;41(3):348–53.

[50] Kato N, Takeuchi F, Tabara Y, Kelly TN, Go MJ, Sim X, et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in East Asians. Nat Genet 2011;43(6):531–8.

[51] Zhu X, Young JH, Fox E, Keating BJ, Franceschini N, Kang S, et al. Combined admixture mapping and association analysis identifies a novel blood pressure genetic locus on 5p13: contributions from the CARe consortium. Hum Mol Genet 2011;20(11):2285–95.

[52] Padmanabhan S, Melander O, Johnson T, Di Blasio AM, Lee WK, Gentilini D, et al. Genome-wide association study of blood pressure extremes identifies variant near UMOD associated with hypertension. PLoS Genet 2010;6(10):e1001177.

[53] Köttgen A, Glazer NL, Dehghan A, Hwang SJ, Katz R, Li M, et al. Multiple loci associated with indices of renal function and chronic kidney disease. Nat Genet 2009;41(6):712–7.

[54] Welander J, Söderkvist P, Gimm O. Genetics and clinical characteristics of hereditary pheochromocytomas and paragangliomas. Endocr Relat Cancer 2011;18(6):R253–76.

[55] van Hulsteijn LT, Dekkers OM, Hes FJ, Smit JW, Corssmit EP. Risk of malignant paraganglioma in SDHB-mutation and SDHD-mutation carriers: a systematic review and meta-analysis. J Med Genet 2012;49(12):768–76.

[56] Touyz RM. Intracellular mechanisms involved in vascular remodelling of resistance arteries in hypertension: role of angiotensin II. Exp Physiol 2005;90(4):449–55.

[57] Xiang M, Feng M, Muend S, Rao R. A human Na+/H+ antiporter sharing evolutionary origins with bacterial NhaA may be a candidate gene for essential hypertension. Proc Natl Acad Sci USA 2007;104(47):18677–81.

[58] Blaustein MP. Physiological effects of endogenous ouabain: control of intracellular Ca2+ stores and cell responsiveness. Am J Physiol 1993;264(6 Pt 1):C1367–87.

[59] Iwamoto T, Kita S, Zhang J, Blaustein MP, Arai Y, Yoshida S, et al. Salt-sensitive hypertension is triggered by Ca2+ entry via Na+/Ca2+ exchanger type-1 in vascular smooth muscle. Nat Med 2004;10(11):1193–9.

[60] Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, et al. Heart disease and stroke statistics–2013 update: a report from the American heart association. Circulation 2013;127(1):e6–e245.

[61] Neal B, MacMahon S, Chapman NCollaboration BPLTT. Effects of ACE inhibitors, calcium antagonists, and other blood-pressure-lowering drugs: results of prospectively designed overviews of randomised trials. Blood Pressure Lowering Treatment Trialists' Collaboration. Lancet 2000;356(9246):1955–64.

[62] Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo Jr JL, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. JAMA 2003;289(19):2560–72.

[63] Perk J, De Backer G, Gohlke H, Graham I, Reiner Z, Verschuren M, et al. European guidelines on cardiovascular disease prevention in clinical practice (version 2012). The fifth joint task force of the European society of cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of nine societies and by invited experts). Eur Heart J 2012;33(13):1635–701.

[64] Liu LS. Hypertension WGoCGftMo. [2010 Chinese guidelines for the management of hypertension]. Zhonghua Xin Xue Guan Bing Za Zhi 2011;39(7):579–615.

[65] Ogihara T, Kikuchi K, Matsuoka H, Fujita T, Higaki J, Horiuchi M, et al. The Japanese society of hypertension guidelines for the management of hypertension (JSH 2009). Hypertens Res 2009;32(1):3–107.

[66] Bramlage P, Böhm M, Volpe M, Khan BV, Paar WD, Tebbe U, et al. A global perspective on blood pressure treatment and control in a referred cohort of hypertensive patients. J Clin Hypertens (Greenwich) 2010;12(9):666–77.

[67] (CDC) CfDCaP. Vital signs: awareness and treatment of uncontrolled hypertension among adults–United States, 2003–2010. MMWR Morb Mortal Wkly Rep 2012;61:703–9.

[68] Bidiville J, Nussberger J, Waeber G, Porchet M, Waeber B, Brunner HR. Individual responses to converting enzyme inhibitors and calcium antagonists. Hypertension 1988;11(2):166–73.

[69] Freis ED, Materson BJ, Flamenbaum V. Comparison of propranolol or hydrochlorothiazide alone for treatment of hypertension. III. Evaluation of the renin-angiotensin system. Am J Med 1983;74(6):1029–41.

[70] Materson BJ, Reda DJ, Cushman WC, Massie BM, Freis ED, Kochar MS, et al. Single-drug therapy for hypertension in men. A comparison of six antihypertensive agents with placebo. The department of veterans affairs cooperative study group on antihypertensive agents. N Engl J Med 1993;328(13):914–21.

[71] Dahlof B, Sever PS, Poulter NR, Wedel H, Beevers DG, Caulfield M, et al. Prevention of cardiovascular events with an antihypertensive regimen of amlodipine adding perindopril as required versus atenolol adding bendroflumethiazide as required, in the Anglo-Scandinavian cardiac outcomes trial-blood pressure lowering arm (ASCOT-BPLA): a multicentre randomised controlled trial. Lancet 2005;366(9489):895–906.

[72] Furberg CD, Wright JT, Davis BR, Cutler JR, Alderman M, Black H, et al. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). JAMA 2002;288(23):2981–97.

[73] Pepine CJ, Handberg EM, Cooper-DeHoff RM, Marks RG, Kowey P, Messerli FH, et al. A calcium antagonist vs a non-calcium antagonist hypertension treatment strategy for patients with coronary artery disease. The International Verapamil-Trandolapril Study (INVEST): a randomized controlled trial. JAMA 2003;290(21):2805–16.

[74] Hansson L, Hedner T, Lund-Johansen P, Kjeldsen SE, Lindholm LH, Syvertsen JO, et al. Randomised trial of effects of calcium antagonists compared with diuretics and beta-blockers on cardiovascular morbidity and mortality in hypertension: the nordic diltiazem (NORDIL) study. Lancet 2000;356(9227):359–65.

[75] Johnson JA, Boerwinkle E, Zineh I, Chapman AB, Bailey K, Cooper-DeHoff RM, et al. Pharmacogenomics of antihypertensive drugs: rationale and design of the pharmacogenomic evaluation of antihypertensive responses (PEAR) study. Am Heart J 2009;157(3):442–9.

[76] Chapman AB, Schwartz GL, Boerwinkle E, Turner ST. Predictors of antihypertensive response to a standard dose of hydrochlorothiazide for essential hypertension. Kidney Int 2002;61(3):1047–55.

[77] Hiltunen TP, Suonsyrja T, Hannila-Handelberg T, Paavonen KJ, Miettinen HE, Strandberg T, et al. Predictors of antihypertensive drug responses: initial data from a placebo-controlled, randomized, cross-over study with four antihypertensive drugs (The GENRES Study). Am J Hypertens 2007;20(3):311–8.

[78] Filigheddu F, Argiolas G, Bulla E, Troffa C, Bulla P, Fadda S, et al. Clinical variables, not RAAS polymorphisms, predict blood pressure response to ACE inhibitors in Sardinians. Pharmacogenomics 2008;9(10):1419–27.

[79] Turner ST, Schwartz GL, Chapman AB, Beitelshees AL, Gums JG, Cooper-DeHoff RM, et al. Plasma renin activity predicts blood pressure responses to beta-blocker and thiazide diuretic as monotherapy and add-on therapy for hypertension. Am J Hypertens 2010;23(9):1014–22.

[80] Hiltunen TP, Kontula K. Clinical and molecular approaches to individualize antihypertensive drug therapy. Ann Med 2012;44(Suppl. 1):S23–9.

[81] Duprez DA. Role of the renin-angiotensin-aldosterone system in vascular remodeling and inflammation: a clinical review. J Hypertens 2006;24(6):983–91.

[82] Laragh JH, Baer L, Brunner HR, Buhler FR, Sealey JE, Vaughan ED. Renin, angiotensin and aldosterone system in pathogenesis and management of hypertensive vascular disease. Am J Med 1972;52(5):633–52.

[83] Laragh JH. Vasoconstriction-volume analysis for understanding and treating hypertension: the use of renin and aldosterone profiles. Am J Med 1973;55(3):261–74.

[84] Laragh JH, Letcher RL, Pickering TG. Renin profiling for diagnosis and treatment of hypertension. JAMA 1979;241(2):151–6.

[85] Preston RA, Materson BJ, Reda DJ, Williams DW, Hamburger RJ, Cushman WC, et al. Age-race subgroup compared with renin profile as predictors of blood pressure response to antihypertensive therapy. Department of veterans affairs cooperative study group on antihypertensive agents. JAMA 1998;280(13):1168–72.

[86] Canzanello VJ, Baranco-Pryor E, Rahbari-Oskoui F, Schwartz GL, Boerwinkle E, Turner ST, et al. Predictors of blood pressure response to the angiotensin receptor blocker candesartan in essential hypertension. Am J Hypertens 2008;21(1):61–6.

[87] Suonsyrjä T, Hannila-Handelberg T, Paavonen KJ, Miettinen HE, Donner K, Strandberg T, et al. Laboratory tests as predictors of the antihypertensive effects of amlodipine, bisoprolol, hydrochlorothiazide and losartan in men: results from the randomized, double-blind, crossover GENRES Study. J Hypertens 2008;26(6):1250–6.

[88] Schwartz GL, Bailey K, Chapman AB, Boerwinkle E, Turner ST. The role of plasma renin activity, age, and race in selecting effective initial drug therapy for hypertension. Am J Hypertens 2013;26(8):957–64.

[89] Brown MJ, Cruickshank JK, Dominiczak AF, MacGregor GA, Poulter NR, Russell GI, et al. Better blood pressure control: how to combine drugs. J Hum Hypertens 2003;17(2):81–6.

[90] Johnson JA, Gong Y, Bailey KR, Cooper-DeHoff RM, Chapman AB, Turner ST, et al. Hydrochlorothiazide and atenolol combination antihypertensive therapy: effects of drug initiation order. Clin Pharmacol Ther 2009;86(5):533–9.

[91] Fan X, Han Y, Sun K, Wang Y, Xin Y, Bai Y, et al. Sex differences in blood pressure response to antihypertensive therapy in Chinese patients with hypertension. Ann Pharmacother 2008;42(12):1772–81.

[92] Turner ST, Bailey KR, Fridley BL, Chapman AB, Schwartz GL, Chai HS, et al. Genomic association analysis suggests chromosome 12 locus influencing antihypertensive response to thiazide diuretic. Hypertension 2008;52(2):359–65.

[93] Messerli FH, Bangalore S. Half a century of hydrochlorothiazide: facts, fads, fiction, and follies. Am J Med 2011;124(10):896–9.

[94] Duarte JD, Cooper-DeHoff RM. Mechanisms for blood pressure lowering and metabolic effects of thiazide and thiazide-like diuretics. Expert Rev Cardiovasc Ther 2010;8(6):793–802.

[95] Cusi D, Barlassina C, Azzani T, Casari G, Citterio L, Devoto M, et al. Polymorphisms of alpha-adducin and salt sensitivity in patients with essential hypertension. Lancet 1997;349(9062):1353–7.

[96] Glorioso N, Filigheddu F, Cusi D, Troffa C, Conti M, Natalizio M, et al. Alpha-adducin 460Trp allele is associated with erythrocyte Na transport rate in North Sardinian primary hypertensives. Hypertension 2002;39(2 Pt 2):357–62.

[97] Glorioso N, Manunta P, Filigheddu F, Troffa C, Stella P, Barlassina C, et al. The role of alpha-adducin polymorphism in blood pressure and sodium handling regulation may not be excluded by a negative association study. Hypertension 1999;34(4 Pt 1):649–54.

[98] Sciarrone MT, Stella P, Barlassina C, Manunta P, Lanzani C, Bianchi G, et al. ACE and alpha-adducin polymorphism as markers of individual response to diuretic therapy. Hypertension 2003;41(3):398–403.

[99] Turner ST, Chapman AB, Schwartz GL, Boerwinkle E. Effects of endothelial nitric oxide synthase, alpha-adducin, and other candidate gene polymorphisms on blood pressure response to hydrochlorothiazide. Am J Hypertens 2003;16(10):834–9.

[100] Turner ST, Boerwinkle E, O'Connell JR, Bailey KR, Gong Y, Chapman AB, et al. Genomic association analysis of common variants influencing antihypertensive response to hydrochlorothiazide. Hypertension 2013;62(2):391–7.

[101] Matayoshi T, Kamide K, Takiuchi S, Yoshii M, Miwa Y, Takami Y, et al. The thiazide-sensitive Na(+)-Cl(−) cotransporter gene, C1784T, and adrenergic receptor-beta3 gene, T727C, may be gene polymorphisms susceptible to the antihypertensive effect of thiazide diuretics. Hypertens Res 2004;27(11):821–33.

[102] Schelleman H, Klungel OH, Witteman JC, Hofman A, van Duijn CM, de Boer A, et al. The influence of the alpha-adducin G460W polymorphism and angiotensinogen M235T polymorphism on antihypertensive medication and blood pressure. Eur J Hum Genet 2006;14(7):860–6.

[103] Schunkert H, Hense HW, Döring A, Riegger GA, Siffert W. Association between a polymorphism in the G protein beta3 subunit gene and lower renin and elevated diastolic blood pressure levels. Hypertension 1998;32(3):510–3.

[104] Turner ST, Schwartz GL, Chapman AB, Boerwinkle E. C825T polymorphism of the G protein beta(3)-subunit and antihypertensive response to a thiazide diuretic. Hypertension 2001;37(2 Part 2):739–43.

[105] Schelleman H, Stricker BH, Verschuren WM, de Boer A, Kroon AA, de Leeuw PW, et al. Interactions between five candidate genes and antihypertensive drug therapy on blood pressure. Pharmacogenomics J 2006;6(1):22–6.

[106] Russo CJ, Melista E, Cui J, DeStefano AL, Bakris GL, Manolis AJ, et al. Association of NEDD4L ubiquitin ligase with essential hypertension. Hypertension 2005;46(3):488–91.

[107] Luo F, Wang Y, Wang X, Sun K, Zhou X, Hui R. A functional variant of NEDD4L is associated with hypertension, antihypertensive response, and orthostatic hypotension. Hypertension 2009;54(4):796–801.

[108] Dahlberg J, Nilsson LO, von Wowern F, Melander O. Polymorphism in NEDD4L is associated with increased salt sensitivity, reduced levels of P-renin and increased levels of Nt-proANP. PLoS One 2007;2(5):e432.

[109] Svensson-Farbom P, Wahlstrand B, Almgren P, Dahlberg J, Fava C, Kjeldsen S, et al. A functional variant of the NEDD4L gene is associated with beneficial treatment response with beta-blockers and diuretics in hypertensive patients. J Hypertens 2011;29(2):388–95.

[110] McDonough CW, Burbage SE, Duarte JD, Gong Y, Langaee TY, Turner ST, et al. Association of variants in NEDD4L with blood pressure response and adverse cardiovascular outcomes in hypertensive patients treated with thiazide diuretics. J Hypertens 2013;31(4):698–704.

[111] Duarte JD, Turner ST, Tran B, Chapman AB, Bailey KR, Gong Y, et al. Association of chromosome 12 locus with antihypertensive

response to hydrochlorothiazide may involve differential YEATS4 expression. Pharmacogenomics J 2013;13(3):257–63.

[112] Mason DA, Moore JD, Green SA, Liggett SB. A gain-of-function polymorphism in a G-protein coupling domain of the human beta1-adrenergic receptor. J Biol Chem 1999;274(18):12670–4.

[113] Johnson JA, Liggett SB. Cardiovascular pharmacogenomics of adrenergic receptor signaling: clinical implications and future directions. Clin Pharmacol Ther 2011;89(3):366–78.

[114] Rathz DA, Brown KM, Kramer LA, Liggett SB. Amino acid 49 polymorphisms of the human beta1-adrenergic receptor affect agonist-promoted trafficking. J Cardiovasc Pharmacol 2002;39(2):155–60.

[115] Levin MC, Marullo S, Muntaner O, Andersson B, Magnusson Y. The myocardium-protective Gly-49 variant of the beta 1-adrenergic receptor exhibits constitutive activity and increased desensitization and down-regulation. J Biol Chem 2002;277(34):30429–35.

[116] Johnson JA, Zineh I, Puckett BJ, McGorray SP, Yarandi HN, Pauly DF. Beta 1-adrenergic receptor polymorphisms and antihypertensive response to metoprolol. Clin Pharmacol Ther 2003;74(1):44–52.

[117] Liu J, Liu ZQ, Yu BN, Xu FH, Mo W, Zhou G, et al. Beta1-Adrenergic receptor polymorphisms influence the response to metoprolol monotherapy in patients with essential hypertension. Clin Pharmacol Ther 2006;80(1):23–32.

[118] Shin J, Johnson JA. Pharmacogenetics of beta-blockers. Pharmacotherapy 2007;27(6):874–87.

[119] O'Shaughnessy KM, Fu B, Dickerson C, Thurston D, Brown MJ. The gain-of-function G389R variant of the beta1-adrenoceptor does not influence blood pressure or heart rate response to beta-blockade in hypertensive subjects. Clin Sci (Lond) 2000;99(3):233–8.

[120] Karlsson J, Lind L, Hallberg P, Michaëlsson K, Kurland L, Kahan T, et al. Beta1-adrenergic receptor gene polymorphisms and response to beta1-adrenergic receptor blockade in patients with essential hypertension. Clin Cardiol 2004;27(6):347–50.

[121] Johnson AD, Newton-Cheh C, Chasman DI, Ehret GB, Johnson T, Rose L, et al. Association of hypertension drug target genes with blood pressure and hypertension in 86,588 individuals. Hypertension 2011;57(5):903–10.

[122] Bhatnagar V, O'Connor DT, Brophy VH, Schork NJ, Richard E, Salem RM, et al. G-protein-coupled receptor kinase 4 polymorphisms and blood pressure response to metoprolol among African Americans: sex-specificity and interactions. Am J Hypertens 2009;22(3):332–8.

[123] Vandell AG, Lobmeyer MT, Gawronski BE, Langaee TY, Gong Y, Gums JG, et al. G protein receptor kinase 4 polymorphisms: β-blocker pharmacogenetics and treatment-related outcomes in hypertension. Hypertension 2012;60(4):957–64.

[124] Beitelshees AL, Navare H, Wang D, Gong Y, Wessel J, Moss JI, et al. CACNA1C gene polymorphisms, cardiovascular disease outcomes, and treatment response. Circ Cardiovasc Genet 2009;2(4):362–70.

[125] Niu Y, Gong Y, Langaee TY, Davis HM, Elewa H, Beitelshees AL, et al. Genetic variation in the beta2 subunit of the voltage-gated calcium channel and pharmacogenetic association with adverse cardiovascular outcomes in the INternational VErapamil SR-Trandolapril STudy GENEtic Substudy (INVEST-GENES). Circ Cardiovasc Genet 2010;3(6):548–55.

[126] Brenner R, Perez GJ, Bonev AD, Eckman DM, Kosek JC, Wiler SW, et al. Vasoregulation by the beta1 subunit of the calcium-activated potassium channel. Nature 2000;407(6806):870–6.

[127] Pluger S, Faulhaber J, Furstenau M, Lohn M, Waldschutz R, Gollasch M, et al. Mice with disrupted BK channel beta1 subunit gene feature abnormal Ca(2+) spark/STOC coupling and elevated blood pressure. Circ Res 2000;87(11):E53–60.

[128] Fernandez-Fernandez JM, Tomas M, Vazquez E, Orio P, Latorre R, Senti M, et al. Gain-of-function mutation in the KCNMB1 potassium channel subunit is associated with low prevalence of diastolic hypertension. J Clin Invest 2004;113(7):1032–9.

[129] Senti M, Fernandez-Fernandez JM, Tomas M, Vazquez E, Elosua R, Marrugat J, et al. Protective effect of the KCNMB1 E65K genetic polymorphism against diastolic hypertension in aging women and its relevance to cardiovascular risk. Circ Res 2005;97(12):1360–5.

[130] Gollasch M, Tank J, Luft FC, Jordan J, Maass P, Krasko C, et al. The BK channel beta1 subunit gene is associated with human baroreflex and blood pressure regulation. J Hypertens 2002;20(5):927–33.

[131] Beitelshees AL, Gong Y, Wang D, Schork NJ, Cooper-Dehoff RM, Langaee TY, et al. KCNMB1 genotype influences response to verapamil SR and adverse outcomes in the INternational VErapamil SR/Trandolapril STudy (INVEST). Pharmacogenet Genomics 2007;17(9):719–29.

[132] Rigat B, Hubert C, Alhenc-Gelas F, Cambien F, Corvol P, Soubrier F. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. J Clin Invest 1990;86(4):1343–6.

[133] Todd GP, Chadwick IG, Higgins KS, Yeo WW, Jackson PR, Ramsay LE. Relation between changes in blood pressure and serum ACE activity after a single dose of enalapril and ACE genotype in healthy subjects. Br J Clin Pharmacol 1995;39(2):131–4.

[134] Nakano Y, Oshima T, Watanabe M, Matsuura H, Kajiyama G, Kambe M. Angiotensin I-converting enzyme gene polymorphism and acute response to captopril in essential hypertension. Am J Hypertens 1997;10(9 Pt 1):1064–8.

[135] Mondorf UF, Russ A, Wiesemann A, Herrero M, Oremek G, Lenz T. Contribution of angiotensin I converting enzyme gene polymorphism and angiotensinogen gene polymorphism to blood pressure regulation in essential hypertension. Am J Hypertens 1998;11(2):174–83.

[136] Harrap SB, Tzourio C, Cambien F, Poirier O, Raoux S, Chalmers J, et al. The ACE gene I/D polymorphism is not associated with the blood pressure and cardiovascular benefits of ACE inhibition. Hypertension 2003;42(3):297–303.

[137] Yu H, Zhang Y, Liu G. Relationship between polymorphism of the angiotensin-converting enzyme gene and the response to angiotensin-converting enzyme inhibition in hypertensive patients. Hypertens Res 2003;26(11):881–6.

[138] Redon J, Luque-Otero M, Martell N, Chaves FJ, Investigators P. Renin-angiotensin system gene polymorphisms: relationship with blood pressure and microalbuminuria in telmisartan-treated hypertensive patients. Pharmacogenomics J 2005;5(1):14–20.

[139] Nordestgaard BG, Kontula K, Benn M, Dahlof B, de Faire U, Edelman JM, et al. Effect of ACE insertion/deletion and 12 other polymorphisms on clinical outcomes and response to treatment in the LIFE study. Pharmacogenet Genomics 2010;20(2):77–85.

[140] Konoshita T, Kato N, Fuchs S, Mizuno S, Aoyama C, Motomura M, et al. Genetic variant of the renin-angiotensin system and diabetes influences blood pressure response to angiotensin receptor blockers. Diabetes Care 2009;32(8):1485–90.

[141] Haas M, Yilmaz N, Schmidt A, Neyer U, Arneitz K, Stummvoll HK, et al. Angiotensin-converting enzyme gene polymorphism determines the antiproteinuric and systemic hemodynamic effect of enalapril in patients with proteinuric renal disease. Austrian study group of the effects of enalapril treatment in proteinuric renal disease. Kidney Blood Press Res 1998;21(1):66–9.

[142] O'Toole L, Stewart M, Padfield P, Channer K. Effect of the insertion/deletion polymorphism of the angiotensin-converting enzyme gene

on response to angiotensin-converting enzyme inhibitors in patients with heart failure. J Cardiovasc Pharmacol 1998;32(6):988–94.

[143] Stavroulakis GA, Makris TK, Krespi PG, Hatzizacharias AN, Gialeraki AE, Anastasiadis G, et al. Predicting response to chronic antihypertensive treatment with fosinopril: the role of angiotensin-converting enzyme gene polymorphism. Cardiovasc Drugs Ther 2000;14(4):427–32.

[144] Suwelack B, Kempkes-Koch M, Kobelt V, Hillebrand U, Matzkies F, Gerhardt U, et al. Impact of ACE polymorphism on renal allograft function, blood pressure, and proteinuria under ACE inhibition. Transplant Proc 2002;34(5):1763–6.

[145] Li X, Du Y, Huang X. Correlation of angiotensin-converting enzyme gene polymorphism with effect of antihypertensive therapy by angiotensin-converting enzyme inhibitor. J Cardiovasc Pharmacol Ther 2003;8(1):25–30.

[146] Jeunemaitre X, Soubrier F, Kotelevtsev YV, Lifton RP, Williams CS, Charru A, et al. Molecular basis of human hypertension: role of angiotensinogen. Cell 1992;71(1):169–80.

[147] Hingorani AD, Jia H, Stevens PA, Hopper R, Dickerson JE, Brown MJ. Renin-angiotensin system gene polymorphisms influence blood pressure and the response to angiotensin converting enzyme inhibition. J Hypertens 1995;13(12 Pt 2):1602–9.

[148] Srivastava K, Chandra S, Bhatia J, Narang R, Saluja D. Association of angiotensinogen (M235T) gene polymorphism with blood pressure lowering response to angiotensin converting enzyme inhibitor (Enalapril). J Pharm Pharm Sci 2012;15(3):399–406.

[149] Brunner M, Cooper-DeHoff RM, Gong Y, Karnes JH, Langaee TY, Pepine CJ, et al. Factors influencing blood pressure response to trandolapril add-on therapy in patients taking verapamil SR (from the International Verapamil SR/Trandolapril [INVEST] Study). Am J Cardiol 2007;99(11):1549–54.

[150] Yu H, Lin S, Jin L, Yu Y, Zhong J, Zhang Y, et al. Adenine/cytosine(1166) polymorphism of the angiotensin II type 1 receptor gene and the antihypertensive response to angiotensin-converting enzyme inhibitors. J Hypertens 2009;27(11):2278–82.

[151] Turner ST, Bailey KR, Schwartz GL, Chapman AB, Chai HS, Boerwinkle E. Genomic association analysis identifies multiple loci influencing antihypertensive response to an angiotensin II receptor blocker. Hypertension 2012;59(6):1204–11.

[152] Dahlof B, Devereux RB, Kjeldsen SE, Julius S, Beevers G, de Faire U, et al. Cardiovascular morbidity and mortality in the Losartan Intervention For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol. Lancet 2002;359(9311):995–1003.

[153] Pacanowski MA, Gong Y, Cooper-Dehoff RM, Schork NJ, Shriver MD, Langaee TY, et al. Beta-adrenergic receptor gene polymorphisms and beta-blocker treatment outcomes in hypertension. Clin Pharmacol Ther 2008;84(6):715–21.

[154] Lemaitre RN, Heckbert SR, Sotoodehnia N, Bis JC, Smith NL, Marciante KD, et al. Beta1- and beta2-adrenergic receptor gene variation, beta-blocker use and risk of myocardial infarction and stroke. Am J Hypertens 2008;21(3):290–6.

[155] Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, et al. Genome-wide association study of blood pressure and hypertension. Nat Genet 2009;41(6):677–87.

[156] Boczek NJ, Best JM, Tester DJ, Giudicessi JR, Middha S, Evans JM, et al. Exome sequencing and systems biology converge to identify novel mutations in the L-type calcium channel, CACNA1C, linked to autosomal dominant long qt syndrome. Circ Cardiovasc Genet 2013;6(3):279–89.

[157] Danesh J, Lewington S, Thompson SG, Lowe GD, Collins R, Kostis JB, et al. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. JAMA 2005;294(14):1799–809.

[158] van 't Hooft FM, von Bahr SJ, Silveira A, Iliadou A, Eriksson P, Hamsten A. Two common, functional polymorphisms in the promoter region of the beta-fibrinogen gene contribute to regulation of plasma fibrinogen concentration. Arterioscler Thromb Vasc Biol 1999;19(12):3063–70.

[159] Fogari R, Zoppi A, Malamani GD, Marasi G, Vanasia A, Villa G. Effects of different antihypertensive drugs on plasma fibrinogen in hypertensive patients. Br J Clin Pharmacol 1995;39(5):471–6.

[160] Lynch AI, Boerwinkle E, Davis BR, et al. Antihypertensive pharmacogenetic effect of fibrinogen-beta variant −455G>A on cardiovascular disease, end-stage renal disease, and mortality: the GenHAT study. Pharmacogenet Genomics 2009;19(6):415–21.

[161] McDonough CW, Gong Y, Padmanabhan S, Burkley B, Langaee TY, Melander O, et al. Pharmacogenomic Association of Nonsynonymous SNPs in SIGLEC12, A1BG, and the selectin region and cardiovascular outcomes. Hypertension 2013;62(1):48–54.

[162] Yu Z, Lai CM, Maoui M, Banville D, Shen SH. Identification and characterization of S2V, a novel putative siglec that contains two V set Ig-like domains and recruits protein-tyrosine phosphatases SHPs. J Biol Chem 2001;276(26):23816–24.

[163] Munday J, Floyd H, Crocker PR. Sialic acid binding receptors (siglecs) expressed by macrophages. J Leukoc Biol 1999;66(5):705–11.

[164] Udby L, Sorensen OE, Pass J, Johnsen AH, Behrendt N, Borregaard N, et al. Cysteine-rich secretory protein 3 is a ligand of alpha1B-glycoprotein in human plasma. Biochemistry 2004;43(40):12877–86.

[165] Mann KG, Kalafatis M. Factor V: a combination of Dr Jekyll and Mr Hyde. Blood 2003;101(1):20–30.

[166] Kansas GS. Selectins and their ligands: current concepts and controversies. Blood 1996;88(9):3259–87.

[167] Auro K, Alanne M, Kristiansson K, Silander K, Kuulasmaa K, Salomaa V, et al. Combined effects of thrombosis pathway gene variants predict cardiovascular events. PLoS Genet 2007;3(7):e120.

[168] Barbaux SC, Blankenberg S, Rupprecht HJ, Francomme C, Bickel C, Hafner G, et al. Association between P-selectin gene polymorphisms and soluble P-selectin levels and their relation to coronary artery disease. Arterioscler Thromb Vasc Biol 2001;21(10):1668–73.

[169] Wenzel K, Felix S, Kleber FX, Brachold R, Menke T, Schattke S, et al. E-selectin polymorphism and atherosclerosis: an association study. Hum Mol Genet 1994;3(11):1935–7.

[170] Psaty BM, Smith NL, Heckbert SR, Vos HL, Lemaitre RN, Reiner AP, et al. Diuretic therapy, the alpha-adducin gene variant, and the risk of myocardial infarction or stroke in persons with treated hypertension. JAMA 2002;287(13):1680–9.

[171] Davis BR, Arnett DK, Boerwinkle E, Ford CE, Leiendecker-Foster C, Miller MB, et al. Antihypertensive therapy, the alpha-adducin polymorphism, and cardiovascular disease in high-risk hypertensive persons: the genetics of hypertension-associated treatment study. Pharmacogenomics J 2007;7(2):112–22.

[172] Gerhard T, Gong Y, Beitelshees AL, Mao X, Lobmeyer MT, Cooper-DeHoff RM, et al. Alpha-adducin polymorphism associated with increased risk of adverse cardiovascular outcomes: results from GENEtic Substudy of the INternational VErapamil SR-trandolapril STudy (INVEST-GENES). Am Heart J 2008;156(2):397–404.

[173] van Wieren-de Wijer DB, Maitland-van der Zee AH, de Boer A, Kroon AA, de Leeuw PW, Schiffers P, et al. Interaction between the Gly460Trp alpha-adducin gene variant and diuretics on the risk of myocardial infarction. J Hypertens 2009;27(1):61–8.

[174] Arnett DK, Davis BR, Ford CE, Boerwinkle E, Leiendecker-Foster C, Miller MB, et al. Pharmacogenetic association of the

angiotensin-converting enzyme insertion/deletion polymorphism on blood pressure and cardiovascular risk in relation to antihypertensive treatment: the genetics of hypertension-associated treatment (GenHAT) study. Circulation 2005;111(25):3374–83.

[175] Schelleman H, Klungel OH, Witteman JC, Breteler MM, Hofman A, van Duijn CM, et al. Interaction between polymorphisms in the renin-angiotensin-system and angiotensin-converting enzyme inhibitor or beta-blocker use and the risk of myocardial infarction and stroke. Pharmacogenomics J 2008;8(6):400–7.

[176] Maitland-van der Zee AH, van Wieren-de Wijer DB, de Boer A, Kroon AA, de Leeuw PW, Schiffers P, et al. Genetic variation in the renin–angiotensin system, use of renin–angiotensin system inhibitors and the risk of myocardial infarction. J Renin Angiotensin Aldosterone Syst 2011;12(3):208–14.

[177] Lynch AI, Eckfeldt JH, Davis BR, Ford CE, Boerwinkle E, Leiendecker-Foster C, et al. Gene panels to help identify subgroups at high and low risk of coronary heart disease among those randomized to antihypertensive treatment: the GenHAT study. Pharmacogenet Genomics 2012;22(5):355–66.

[178] Morello F, de Boer RA, Steffensen KR, Gnecchi M, Chisholm JW, Boomsma F, et al. Liver X receptors alpha and beta regulate renin expression *in vivo*. J Clin Invest 2005;115(7):1913–22.

[179] Price ET, Pacanowski MA, Martin MA, Cooper-DeHoff RM, Pepine CJ, Zineh I, et al. Liver X receptor alpha gene polymorphisms and variable cardiovascular outcomes in patients treated with antihypertensive therapy: results from the INVEST-GENES study. Pharmacogenet Genomics 2011;21(6):333–40.

[180] Tanner RM, Lynch AI, Brophy VH, Eckfeldt JH, Davis BR, Ford CE, et al. Pharmacogenetic associations of MMP9 and MMP12 variants with cardiovascular disease in patients with hypertension. PLoS One 2011;6(8):e23609.

[181] Colombo MG, Paradossi U, Andreassi MG, Botto N, Manfredi S, Masetti S, et al. Endothelial nitric oxide synthase gene polymorphisms and risk of coronary artery disease. Clin Chem 2003;49(3):389–95.

[182] Hibi K, Ishigami T, Tamura K, Mizushima S, Nyui N, Fujita T, et al. Endothelial nitric oxide synthase gene polymorphism and acute myocardial infarction. Hypertension 1998;32(3):521–6.

[183] Colombo MG, Andreassi MG, Paradossi U, Botto N, Manfredi S, Masetti S, et al. Evidence for association of a common variant of the endothelial nitric oxide synthase gene (Glu298→Asp polymorphism) to the presence, extent, and severity of coronary artery disease. Heart 2002;87(6):525–8.

[184] Lenasi H, Kohlstedt K, Fichtlscherer B, Mulsch A, Busse R, Fleming I. Amlodipine activates the endothelial nitric oxide synthase by altering phosphorylation on Ser1177 and Thr495. Cardiovasc Res 2003;59(4):844–53.

[185] Zhang X, Lynch AI, Davis BR, Ford CE, Boerwinkle E, Eckfeldt JH, et al. Pharmacogenetic association of NOS3 variants with cardiovascular disease in patients with hypertension: the GenHAT study. PLoS One 2012;7(3):e34217.

[186] Ackermann U. Structure and function of atrial natriuretic peptides. Clin Chem 1986;32(2):241–7.

[187] Rubattu S, Stanzione R, Di Angelantonio E, Zanda B, Evangelista A, Tarasi D, et al. Atrial natriuretic peptide gene polymorphisms and risk of ischemic stroke in humans. Stroke 2004;35(4):814–8.

[188] Dedoussis GV, Maumus S, Skoumas J, Choumerianou DM, Pitsavos C, Stefanadis C, et al. Natriuretic peptide Val7Met substitution and risk of coronary artery disease in Greek patients with familial hypercholesterolemia. J Clin Lab Anal 2006;20(3):98–104.

[189] Lynch AI, Boerwinkle E, Davis BR, Ford CE, Eckfeldt JH, Leiendecker-Foster C, et al. Pharmacogenetic association of the NPPA T2238C genetic variant with cardiovascular disease outcomes in patients with hypertension. JAMA 2008;299(3):296–307.

[190] Kohler R, Brakemeier S, Kuhn M, Degenhardt C, Buhr H, Pries A, et al. Expression of ryanodine receptor type 3 and TRP channels in endothelial cells: comparison of in situ and cultured human endothelial cells. Cardiovasc Res 2001;51(1):160–8.

[191] Shrestha S, Irvin MR, Taylor KD, Wiener HW, Pajewsk NM, Haritunians T, et al. A genome-wide association study of carotid atherosclerosis in HIV-infected men. AIDS 2010;24(4):583–92.

[192] Yano M, Yamamoto T, Ikeda Y, Matsuzaki M. Mechanisms of disease: ryanodine receptor defects in heart failure and fatal arrhythmia. Nat Clin Pract Cardiovasc Med 2006;3(1):43–52.

[193] Shibata MC, Leon H, Chatterley T, Dorgan M, Vandermeer B. Do calcium channel blockers increase the diagnosis of heart failure in patients with hypertension? Am J Cardiol 2010;106(2):228–35.

[194] Lynch AI, Irvin MR, Boerwinkle E, Davis BR, Vaughan LK, Ford CE, et al. RYR3 gene polymorphisms and cardiovascular disease outcomes in the context of antihypertensive treatment. Pharmacogenomics J 2013;13(4):330–4.

[195] Schelleman H, Klungel OH, Witteman JC, Breteler MM, Yazdanpanah M, Danser AH, et al. Angiotensinogen M235T polymorphism and the risk of myocardial infarction and stroke among hypertensive patients on ACE-inhibitors or beta-blockers. Eur J Hum Genet 2007;15(4):478–84.

[196] Lynch AI, Eckfeldt JH, Davis BR, Ford CE, Boerwinkle E, Leiendecker-Foster C, et al. Gene panels to help identify subgroups at high and low risk of coronary heart disease among those randomized to antihypertensive treatment: the GenHAT study. Pharmacogenet Genomics 2012;22(5):355–66.

[197] Gong Y, McDonough CW, Wang Z, Hou W, Cooper-DeHoff RM, Langaee TY, et al. Hypertension susceptibility loci and blood pressure response to antihypertensives: results from the pharmacogenomic evaluation of antihypertensive responses study. Circ Cardiovasc Genet 2012;5(6):686–91.

[198] Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 2008;359(21):2208–19.

[199] Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. Lancet 2010;376(9750):1393–400.

[200] Collins F. Opportunities and challenges for the NIH–an interview with Francis Collins. Interview by Robert Steinbrook. N Engl J Med 2009;361(14):1321–3.

[201] Benavente OR, White CL, Pearce L, Pergola P, Roldan A, Benavente MF, et al. The secondary prevention of small subcortical strokes (SPS3) study. Int J Stroke 2011;6(2):164–75.

[202] Padmanabhan S, Newton-Cheh C, Dominiczak AF. Genetic basis of blood pressure and hypertension. Trends Genet 2012;28(8):397–408.

Chapter 33

QTc and Sudden Cardiac Death

Catriona Brown and Sandosh Padmanabhan
Institute of Cardiovascular and Medical Sciences, Glasgow, UK

Chapter Outline

33.1 Overview 779
33.1.1 The Resting ECG 779
33.1.2 The QT Interval 780
33.1.3 Torsades de Pointes 781
33.1.4 Corrected QT Interval 781
33.1.5 Cardiac Myocyte Ion Currents and Action Potentials 781
33.2 Genetic Syndromes 784
33.2.1 Congenital Long-QT Syndrome 784
33.3 Drug-Induced Variation in QTc 787
33.3.1 The Human Ether-à-go-go Gene Channel 788
33.3.2 Repolarization Heterogeneity 788
33.3.3 Mechanisms of Proarrhythmia in QT Prolongation 788
33.3.4 Reduced Repolarization Reserve 790
33.3.5 QTc Prolongation–Inducing Drugs—A Closer Look 791
33.3.6 Pharmacokinetic and Pharmacodynamic Factor 794
33.3.7 Prevention, Monitoring and Treatment of Drug-Induced LQTS 795
33.3.8 Detecting Drug-Induced Effects on Repolarization 795
33.3.9 Congenital LQTS Genes in Drug-Induced LQTS 795
33.3.10 *KCNH2* Polymorphisms and Disease Susceptibility 796
33.4 Public Health Impact 796
33.5 QTc in New Drug Discovery and Drug Safety 797
33.6 Pharmacogenomic Studies 798
33.6.1 HERG Potassium Channels, *KCNH2* Polymorphisms, and Drug-Induced LQTS 798
33.6.2 Functional Deficits Caused by HERG Channel Mutations 799
33.6.3 Distribution of Drug Sensitivity–Related *KCNH2* Polymorphisms 799
33.6.4 Rare Mutations 800
33.6.5 Common Variants and Common Polymorphisms in Ion Channel Genes 800
33.6.6 Genome-Wide Association Studies 800
33.7 Future Directions 801
33.7.1 Screening, Stratification, Personalization 801
33.7.2 Proposed Methods of Risk Stratification 802
33.7.3 Unanswered Questions and Priority Areas for Research 802
Glossary 802
Acronyms and Abbreviations 803
References 803
Suggested websites 806

33.1 OVERVIEW

33.1.1 The Resting ECG

Muscles contract as a result of electrical changes known as "depolarizations." Such changes can be detected by electrodes attached on the surface of the body [1]. In order to measure depolarizations in heart muscle during a resting ECG, the patient must be fully relaxed, with no other skeletal muscles contracting. Standard readings are in the form of a 12-lead electrocardiogram (ECG), the 12-lead tracing being derived from six chest and four limb electrodes. Before going further, it is important to briefly summarize the anatomy of the heart. Blood enters the right atrium from the rest of the body via two large veins: the inferior vena cava from below and the superior vena cava from above. The left atrium receives blood from the lungs via the pulmonary veins. The atria are the two upper chambers of the heart, and their contraction forces blood into the lower two

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00033-5
© 2014 Elsevier Inc. All rights reserved.

chambers, the ventricles. The blood is then pumped out of the heart by the ventricles to supply the lungs (from the right ventricle via the pulmonary artery) and the rest of the body (from the left ventricle via the aorta). Depolarization begins in the sinoatrial (SA) node and spreads as a wave through the atrial muscles. There is then delay as depolarization spreads through the atrioventricular (AV) node, but thereafter conduction is rapid as it travels down specialized conduction tissue [1]. This begins as a single pathway, the "bundle of His," and then divides into right and left branches. Within the mass of ventricular muscle, conduction spreads rapidly through "Purkinje fibers" which are specialized tissues. The speeds of conduction through the various tissues are as follows: atrial muscle 1 meter per second (m/s); AV node 0.2 m/s; His bundle, bundle branches, and Purkinje tissue 4.0 m/s; and ventricular muscle 0.5 m/s [1]. On the ECG tracing, the P wave represents depolarization of the atria, (see Figure 33.1). The PR interval represents conduction through the AV node and His bundle and measures the time it takes for the depolarization wave to spread from the atria to the interventricular septum. The QRS complex is the combination of three deflections "Q", "R" and "S" (see Figure 33.1) seen in rapid succession and represents depolarization of the ventricles. The T wave represents ventricular muscle repolarization [1]. U waves can occasionally be seen as extra humps on the end of a T wave and are thought to represent repolarization of the papillary muscles of the heart. However, they can also be indicative of hypokalaemia.

Normally each heartbeat consists of a P wave followed by a QRS complex. The QRS complex should then have only one associated T wave. The time taken for each of these sections of the ECG can be measured. Normal paper speed is 25 millimeters per second (mm/s), one large square represents 200 milliseconds (ms), and one small square represents 40 ms. The normal limits are as follows: PR interval, 120–200 ms; QRS complex, <120 ms; T-wave duration is difficult to measure [1]. Therefore, the QT interval, discussed in the next section, is usually taken as the best indication of the time taken for repolarization of the ventricles.

33.1.2 The QT Interval

On a 12-lead ECG, the QT interval is measured from the beginning of the QRS complex to the end of the T wave as it returns to baseline [2]. It can be affected by both ventricular conduction velocities and the velocity of repolarization. Manual measurements of the QT interval should be taken from leads II and V5 or V6, with the longest value being used. A mean value should be derived from at least 3–4 cardiac cycles. The end of the T wave can be determined reliably by the slope method [2], defined by the intersection point between the tangent drawn at the maximum downslope of the T wave and the isoelectric line. If the T wave is notched, the tangent should be applied to the maximum slope as described by Postema et al. [2]. (See Figure 33.1.) Long-QT syndrome (LQTS) is characterized by a prolonged QT interval in the surface ECG, and can be associated with a propensity to develop syncope and sudden cardiac death (SCD). In most documented cases, death was caused by the malignant polymorphic ventricular arrhythmia called torsades de pointes (TdP). Cardiomyocytes differ from other tissues (neurons and skeletal muscle) in that they possess a long action potential, approximately 300 ms [3]. For normal sinus rhythm to be maintained, an ordered pattern of conduction and repolarization is essential. The normal QT in healthy people without other cardiac conductance–altering factors is usually ~400 msec when measuring this interval in lead II of the ECG. Taking the measurement in this lead is suggested because it is considered to have the longest interval [3]. The absolute value of the measured QT interval should be corrected for heart rate because the QT interval varies with changes in heart rate. For example, in tachycardia the action potential will be shorter, and the absolute measure of QT will also be shorter as a consequence of this [3] (see Section 33.1.4).

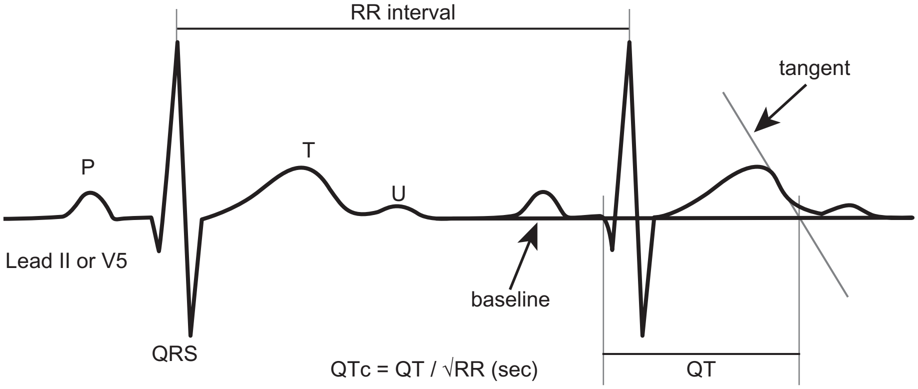


FIGURE 33.1 **QT interval.** A tangent is drawn to the steepest slope of the last limb of the T wave in lead II or V5. The end of the T wave is the intersection of the tangent with the baseline. QT is heart rate corrected with Bazett's formula using the preceding RR interval. *Source: Reprinted from Postema et al.* [2] *with permission.*

33.1.3 Torsades de Pointes

In 1966, Dessertenne reported his index case of an elderly woman with complete AV block and syncopal attacks due to a polymorphic ventricular tachycardia (VT) [4]. He termed this condition *torsades de pointes* (TdP). The definition of TdP is as follows: a polymorphic ventricular tachycardia characterized by a "twisting of the points" around the isoelectric line on the electrocardiogram (ECG) preceded by a long QT interval [3]. The polymorphic nature of TdP may not always be evident, particularly when the arrhythmia is recorded on only one or two leads. A pause-dependent tachycardia associated with severe QT prolongation may therefore appear to be "monomorphic" when in fact it represents TdP [5]. Such diagnostic difficulties reinforce the need for documentation of the index arrhythmia event and for review by expert physicians. It is also important to note that TdP can be seen in both congenital and acquired long QT syndrome.

TdP is often self-limiting, producing transient syncope, but it can also degenerate into ventricular fibrillation and cause cardiac arrest or sudden death. We still do not know why in some patients TdP stops after a few seconds whereas in others it continues with serious or devastating consequences. Although QT prolongation is clearly associated with an increased risk of TdP, there is no reliable criterion to identify the length of QT prolongation that is associated with a clinically significant increased risk of TdP [3]. This makes it difficult for clinicians to decide what QT interval represents significant risk of TdP and requires intervention. In acquired LQTS, torsades de pointes is characteristically preceded by a "short-long-short" ventricular sequence ("pause-dependent" phenomenon) [5]. Common associated features are pause-dependent changes in T-wave morphology and late coupled ventricular ectopy, which is often slower than coarse ventricular fibrillation [5]. See Figure 33.2a for an example of a short-long-short ECG sequence and Figure 33.2b for an example of QRS complexes twisting around an isoelectric line in Torsades de Pointes. Figure 33.2c from Kannankeril et al. [6] shows the 12-lead ECG of an elderly patient with advanced heart disease who had begun taking dofetilide. An abnormal QT interval is followed by a pause, then four beats of polymorphic ventricular tachycardia. After another pause there is sustained TdP.

In some types of congenital LQTS it is postulated that TdP is more likely to be triggered by adrenergic stimulation, and in such cases, is classified as adrenergic-dependent. However, these classifications are not mutually exclusive; adrenergic-dependent TdP is more likely to occur in the presence of pauses or slow heart rates [7].

33.1.4 Corrected QT Interval

It is generally accepted that in the context of drug-induced QT prolongation, an increase in the corrected QT (QTc) interval of >40 ms is alarming, and increases in mean QTc of 6 ms are important. In clinical practice, diagnosis of potentially serious QT interval changes requires this trait to be measured with precision [8,9]. The risk of proarrhythmia varies from drug to drug and may be a patient-specific response. It is unpredictable and influenced by several predisposing factors discussed in Section 33.3. However, QT interval assessment does remain a prerequisite for drug development, approval, and clinical testing [9].

The very real problem of erroneous QTc increases (overcorrection) at heart rates of more than 60 beats/min, or QTc decreases (undercorrection) at heart rates of fewer than 60 beats/min must always be taken into consideration by clinicians and researchers as well as the pharmaceutical industry and regulatory bodies [9]. The role of drug safety assessment and the "thorough QT" test is discussed later in this chapter. Based on analysis of digitized data for QT and RR interval measurements in healthy individuals, and using Bazett's formula, the ranges of QTc in Table 33.1 have been suggested for diagnosis of long-QT syndrome [10]. These values are generally accepted in research and clinical practice.

The various methods of correcting QT interval are summarized as follows:

- Bazett: $QTcB = QT/RR^{1/2}$ [9]
- Fridericia: $QTcFi = QT/RR^{1/3}$ [9]
- Framingham-Sagie: $QTcFa = QT + 154\,(1\text{–}60/\text{heartrate})$ [11]
- Hodges: $QTcH = QT + 1.75\ (\text{heart rate}\text{–}60)$ [12]

Chiladakis et al. showed that the different QTc formulas provide significantly different ΔQTc values at lower as well as higher heart rates, whereby the QTcB exhibits the highest ΔQTc variability across the various heart rate stages investigated (60 beats/min, 80 beats/min, and 100 beats/min). Conversely, Hodge's formula showed the least heart rate–dependent ΔQTc values followed by Fridericia, and Framingham [9]. Despite these findings, Bazett's formula remains the most commonly used method of QTc correction; however, its limitations should be remembered, and in borderline cases it may be useful to perform QTc correction by one of the methods with a ΔQTc that is less heart rate–dependent.

33.1.5 Cardiac Myocyte Ion Currents and Action Potentials

The QT interval of the ECG is predominantly determined by duration of the action potential (AP) of ventricular cells. Indeed, most cases of QT interval prolongation are caused by factors that prolong the duration of the AP, mainly by delaying repolarization phase 3 [3]. The normal synchronous contraction of the atria and of the ventricles in sequence require rapid activation of groups of cardiac cells [13].

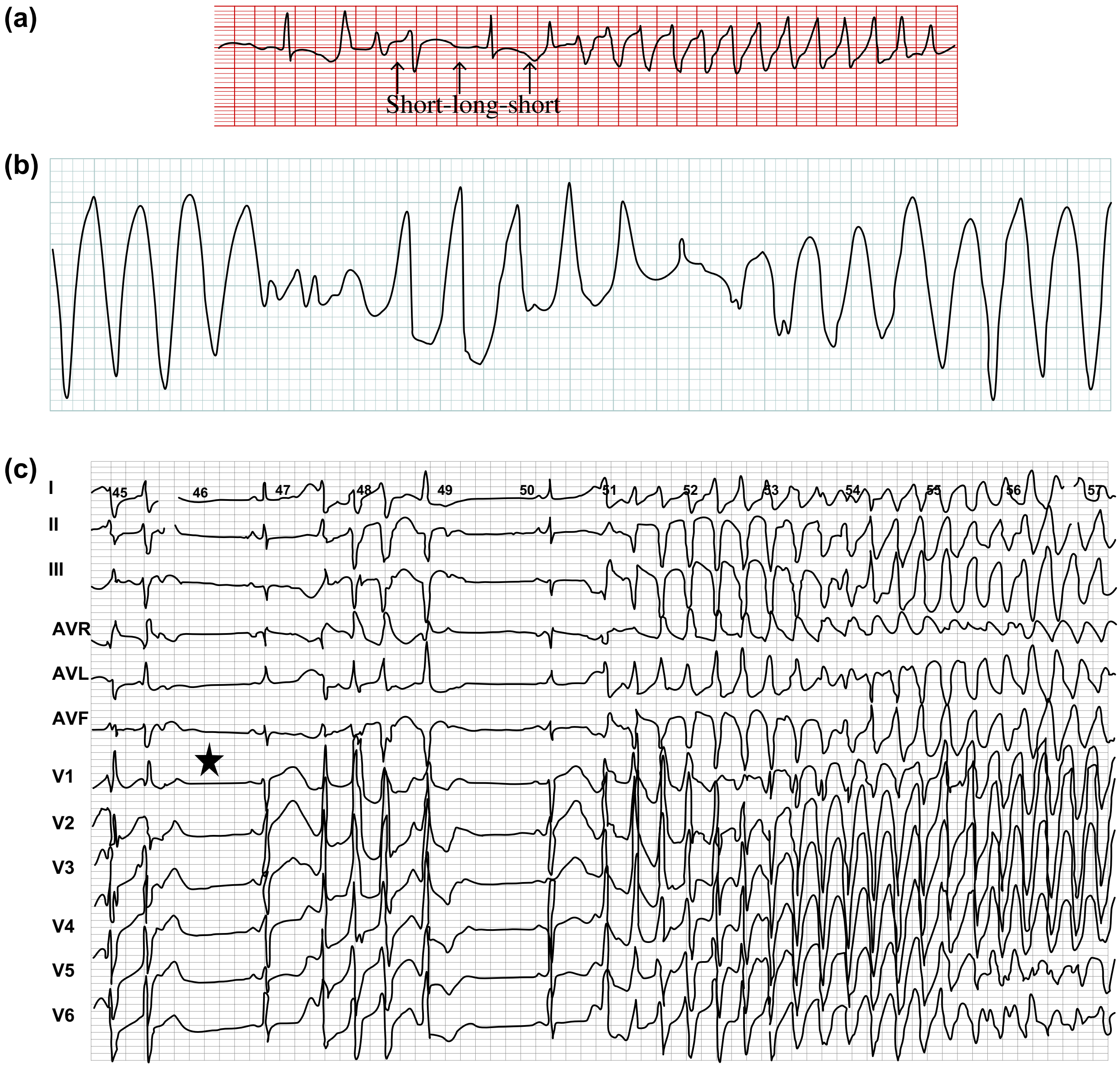


FIGURE 33.2 **Short-long-short sequence preceding TdP; TdP, and drug-induced LQTS.** (a) TdP is usually preceded by a short-long-short ECG sequence. Here one or more premature ventricular complexes are followed by a compensatory pause. The subsequent sinus beat may have an especially long QT and deformities of T or U waves. This sinus beat is followed by another premature ventricular complex that precipitates TdP. (b) The QRS complexes are twisting around an isoelectric line in a sinusoidal fashion. (c) A common feature is a pause (often after an ectopic beat) with deranged repolarization in the following cycle. Shown is a 12-lead ECG recorded from a 79-year-old patient with advanced heart disease who had recently begun taking dofetilide. The abnormal QT interval is followed by a pause (star) and then four beats of polymorphic ventricular tachycardia (TdP). Sustained TdP occurs after another pause. *Source: (a) and (b) from Kallergis et al.* [121]; *(c) from Kannankeril et al.* [6] *with permission*.

An activation mechanism must enable rapid changes in heart rate and also respond to changes in autonomic tone. The propagating cardiac action potential fulfils these roles. More than 20 currents subdivided as depolarizing and repolarizing are involved in the AP generation of ventricular cells [3].

There are five phases in a normal action potential (Figure 33.3). Phase 4, the resting potential, is stable at ~−90 mV in normal working myocardial cells [13]; it is followed by phase 0 (upstroke depolarization), caused by the very rapid activation of voltage-gated sodium channels

TABLE 33.1 Suggested Bazett-Corrected QTc Values for Diagnosing QT Prolongation

Rating	1–15 yrs (ms)	Adult Male (ms)	Adult Female (ms)
Normal	<440	<430	<450
Borderline	440–460	430–450	450–470
Prolonged	>460	>450	>470

ms = milliseconds.
Source: Reprinted with permission from Goldenberg I, Moss AJ, Zareba W. QT Interval: how to measure it and what is "normal." *J Cardiovasc Electrophysiol* 2006;17:333–6.

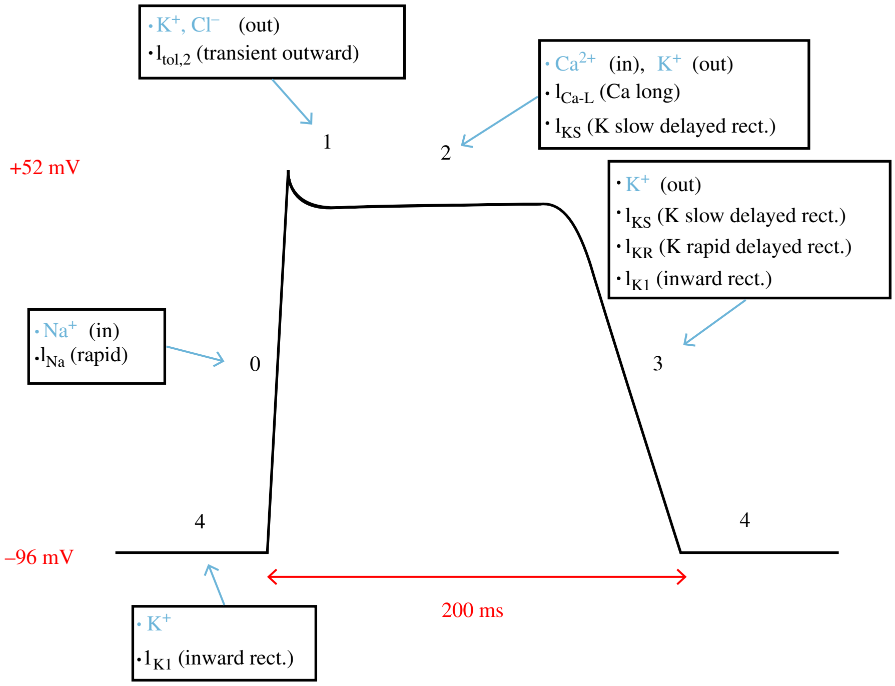

• K^+, Cl^- (out)
• $I_{tol,2}$ (transient outward)
• Ca^{2+} (in), K^+ (out)
• $I_{Ca\text{-}L}$ (Ca long)
• I_{KS} (K slow delayed rect.)
1
2
+52 mV
• K^+ (out)
• I_{KS} (K slow delayed rect.)
• I_{KR} (K rapid delayed rect.)
• I_{K1} (inward rect.)
• Na^+ (in)
• I_{Na} (rapid)
0
3
4
4
–96 mV
200 ms
• K^+
• I_{K1} (inward rect.)

FIGURE 33.3 **Myocardial action potential.** Phase 0 rapid depolarization is mediated by sodium entry into cells. Phase 1 and 3 repolarization results from potassium efflux from cells. Balanced slow calcium entry and potassium exit cause the plateau in phase 2. Potassium re-enters and sodium exits cells during phase 4 recovery. *Source: From Kallergis et al.* [121].

Nav1.5. Next is phase 1, a phase of rapid repolarization due to the transient outward K^+ current. Phase 2, the plateau phase is maintained by inward Ca^{2+} current flowing through slowly inactivating Cav1.2 voltage-gated channels. This is the longest phase. Phase 3 is a phase of rapid depolarization which restores the membrane potential back to its resting value [13]. It occurs through the action of outward currents including: I_{Ks} (slow delayed rectifier potassium current) and I_{Kr} (rapid delayed rectifier potassium current). The main pore-forming subunit (alpha subunit) of I_{Ks} is called KvLQT1 or *KCNQ1* (new nomenclature: Kv7.1); of I_{Kr}, HERG or *KCNH2* (new nomenclature: Kv11.1) [3]. The currents generated by all other ion transporters (ATPases, exchangers, and ion channels) also contribute to shape the AP, as shown in Figure 33.3.

The action potentials of pacemaker cells in the sinoatrial (SA) and atrioventricular (AV) nodes are significantly different from those in the myocardium. In the pacemaker cells, the membrane potential at the onset of phase 4 is more depolarized (−50 to −65 mV). It undergoes slow diastolic depolarization and gradually merges into phase 0 [13]. Indeed, rate of depolarization in phase 0 is much slower than that in the myocardial cells. This ultimately results in slow propagation of the cardiac impulse in the nodal regions [13].

Characteristics of the action potential are also different across the myocardial wall from endocardium, mid-myocardium, to epicardium [13], with epicardial cells having a prominent phase 1 and the shortest action potential. Mid-myocardial cells have the longest action potential duration. Overall, prolonged repolarization in cardiac ventricular myocytes may be due to an increase in depolarizing inward currents or to a decrease in repolarizing outward currents.

A decrease in repolarizing outward currents mostly involves potassium channels, in particular the delayed rectifier potassium current, I_K in cases of drug-induced LQTS [3]. The rapid delayed rectifier potassium current (I_{Kr}) is the primary molecular target for most blocking drugs known to cause TdP. It links congenital and acquired LQTS. Interestingly, it was only through finding genetic mutations in congenital LQTS that the role of specific ion channels in human physiology could be ascertained [3].

Some animal models have investigated drug-induced TdP occurring after pretreatment with the α-blocker methoxamine or after complete atrioventricular nodal ablation [14]. A likely mechanism in both situations is the inhibition of I_{Ks} and enhancement of the pharmacologic effect of the I_{Kr} blockade. Striking action potential lability on drug exposure, in the dog model of atrioventricular nodal ablation, separates arrhythmia-prone from arrhythmia-resistant animals. This suggests that measurement of the QT interval alone may not be sufficient in assessing risk, as other indices of repolarization are also important [14].

33.2 GENETIC SYNDROMES

Congenital and drug-induced long-QT syndromes were beginning to be described in the 1950s; as early as the 1960s, the link between congenital and drug-induced long-QT syndrome was being considered.

33.2.1 Congenital Long-QT Syndrome

Congenital LQTS has a prevalence of 1 in 2000 to 1 in 3000 [15] and may account for up to 13% of sudden infant death [16]. The rare congenital forms of long-QT syndrome (cLQTS) are mainly divided into the autosomal recessive variety associated with deafness (first described by Jervell and Lange-Nielsen in 1957) and the more prevalent autosomal dominant Romano-Ward syndrome (described by Romano et al. in 1963 and Ward in 1964). There are now 13 genes implicated in congenital long-QT syndrome (see Table 33.2) and novel genes continue to be identified.

LQT1, LQT2, and LQT3 phenotypes account for the majority of cLQTS cases. LQT1 accounts for 40–55% [17,18]. The LQT1 gene is *KCNQ1*, isolated on chromosome 11p15.5. *KCNQ1* encodes potassium channel subunits, which are highly expressed in the heart. Four KvLQT1 alpha subunits interact with other proteins (e.g., the minK beta subunit) to create the IK_s ion channel [17]. This channel is responsible for the delayed potassium rectifier current of the cardiac action potential, as mentioned previously. Mutations in the *KCNQ1* gene can be inherited in an autosomal dominant or recessive manner, and it is implicated in Romano-Ward and Jervell and Lange-Nielsen syndromes. The LQT1 phenotype is characterized by events induced by exercise.

LQT2 accounts for 30–35% of cases of congenital LQTS [17,18]. It is caused by a variety of mutations in the *human ether-à-go-go gene (HERG)* (also known commonly as *KCNH2*) See Section 33.3.1. This is also a potassium channel gene, located on chromosome 7. The *KCNH2* gene encodes the rapid component of the potassium rectifying current (IK_r). The mutations may involve the pore or the nonpore region of the HERG channel. Pore mutations carry a high risk for cardiac events and may affect young patients [19], whereas nonpore mutations often lead to TdP in the presence of hypokalemia. The most common trigger in LQT2 tends to be emotional stress [18]. LQT3 accounts for 5–10% of cardiac event cases [17,18]. It is caused by mutations in the sodium channel gene (*SCN5A*), located on chromosome 3p21–24, and is characterized by events occurring at rest or during sleep.

Each of the three major genotypes has a distinctively different T-wave repolarization pattern on ECG, with LQT1 exhibiting a broad-based T-wave pattern, LQT2 exhibiting a notched or bifid T wave, and LQT3 having a typical late-onset peaked or biphasic T wave [10]. Studies investigating triggers for cLQTS have also found that, of patients experiencing cardiac events while swimming, 99% had LQT1; of patients experiencing events triggered by loud noises (e.g., alarm clocks), 80% had LQT2 [18].

Criteria for the diagnosis of congenital long-QT syndrome were first introduced in 1985 by Schwartz [20] and consisted of three major and four minor criteria. The criteria were updated in 1993 (see Table 33.3) [21] to take into account the role of sex-related differences, the wide spectrum of clinical abnormalities observed in LQTS, and the contribution of molecular biology. This set of more specific diagnostic criteria provides a quantitative score to discriminate between patients likely or unlikely to be affected by cLQTS [22].

Now, with at least 13 disease-causing LQTS genes identified, we can become more reliant on genotyping to assist us with diagnosis in cLQTS. However, patients who are positive for a cLQTS disease–causing genotype may be completely asymptomatic, with a normal QTc on 12-lead ECG. This can make diagnosis a challenge and can be explained by the concept of low penetrance. Penetrance represents the probability that a genotype positive individual will manifest clinical signs of the condition. In 1999 Priori et al. investigated penetrance in cLQTS families and found that out of 5 families investigated,

TABLE 33.2 Congenital Long-QT Syndrome Genes

Disease Subtype	OMIM Reference Number (www.omim.org)	Syndrome	Gene	Locus	Frequency in LQTS	Gene Protein Product	Functional Effect
LQT1	#192500	RWS, JLNS	*KCNQ1*	11p15.5	40–55%	I_{Ks} channel, α-subunit	Loss-of-function
LQT2	#613688	RWS	*KCNH2*	7q35-36	30–45%	I_{Kr} channel, α-subunit	Loss-of-function
LQT3	#603830	RWS	*SCN5A*	3p21	5–10%	I_{Na} channel, α-subunit	Gain-of-function
LQT4	#600919	RWS	*ANK2*	4q25-q27	<1%	Ankyrin 2	Loss-of-function
LQT5	#613695	RWS, JLNS	*KCNE1*	21q22.1	<1%	I_{Ks} channel, β-subunit	Loss-of-function
LQT6	#613693	RWS	*KCNE2*	21q22.1	<1%	I_{Kr} channel, β-subunit	Loss-of-function
LQT7	#600681	AS	*KCNJ2*	17q23.1	<1%	I_{Kl} channel, α-subunit	Loss-of-function
LQT8	#601005	TS	*CACNA1C*	12p13.3	<1%	I_{CaL} channel, α-subunit	Gain-of-function
LQT9	#611818	RWS	*CAV3*	3p25.3	<1%	Caveolin-3	Gain of function of Na current
LQT10	#611819	RWS	*SCN4B*	11q23.3	<1%	I_{Na} channel, β-subunit 4	Gain-of-function
LQT11	#611820	RWS	*AKAP9*	7q21-q22	<1%	Yotiao	Reduced IKs current
LQT12	#612955	RWS	*SNTA1*	20q11.2	<1%	α1-syntrophin	Increased Na current
LQT13	#613485	RWS	*KCNJ5*	11q24	<1%	I_{KACh} channel, α-subunit	Loss-of-function

KCNQ1—potassium voltage-gated channel, KQT-like subfamily, member 1; RWS—Romano-Ward syndrome; JLNS—Jervell and Lange-Nielsen syndrome; *KCNH2*—potassium voltage-gated channel, subfamily H, member 2; *SCN5A*—sodium voltage-gated channel, type V, α subunit; *ANK2*—ankyrin 2; *KCNE1*—potassium voltage-gated channel, I_{SK}-related subfamily, member1; *KCNE2*—potassium voltage-gated channel, I_{SK}-related subfamily, member 2; *KCNJ2*—potassium channel, inwardly rectifying, subfamily J, member 2; AS—Andersen syndrome; *CACNA1C*—calcium voltage-dependent channel, L-type, α-1C subunit; TS—Timothy syndrome; *CAV3*—caveolin 3; *SCN4B*—sodium voltage-gated channel, type IV, β subunit; *AKAP9*—A-kinase anchor protein 9; *SNTA1*—syntrophin α1; *KCNJ5*—potassium channel, inwardly rectifying, subfamily J, member 5.

penetrance varied from 14%–33%. When family history of sudden death at a young age was included (with those dying young assumed to be genotype positive), penetrance varied from 25%–60% [23].

Other factors to be considered in the explanation of incomplete penetrance in cLQTS is the presence of multiple mutations. Mullally et al. [24] assessed the risk of life-threatening cardiac events in patients with multiple cLQTS mutations. Mutations (numbering ≥ 2) could involve either; the same allele of one cLQTS-associated gene, both alleles of the same cLQTS-associated gene, or different cLQTS genes. Patients with autosomal recessive Jervell and Lange-Nielsen syndrome were excluded. Patients with multiple mutations were found to have a longer QTc at enrollment, and had a greater risk for life-threatening cardiac events than those with a single cLQTS mutation [24].

The yield of genetic testing has been found to be highest (72%) among those tested having the highest clinical probability [10]. However, a negative genetic test in a subject with clinical cLQTS (i.e., genotype negative/phenotype positive LQTS) provides no basis for excluding the diagnosis. A positive genetic test may be beneficial in influencing treatment decisions and may provide the means for precise “carrier” status classification of potentially at-risk relatives.

TABLE 33.3 Diagnostic Criteria LQTS—1993 Schwartz

ECG Findings[a]	
(A) QTc[b]	
• ≥480 ms	3 points
• 460–479 ms	2 points
• 450–459 ms (in males)	1 point
(B) Torsades de Pointes[c]	2 points
(C) T wave alternans	1 point
(D) Notched T wave in three leads	1 point
(E) Low heart rate for age[d]	0.5 point
Clinical history	
(A) Syncope[c]	
• With stress	2 points
• Without stress	1 point
(B) Congenital deafness	0.5 point
Family history[e]	
(A) Family members with definite LQTS[f]	1 point
(B) Unexplained sudden cardiac death under 30yrs in immediate family members	0.5 point

[a] *In the absence of medications or disorders known to affect these ECG features.* [b] *Calculated by Bazett's formula, where* $QTc = QT/\sqrt{RR}$. [c] *Mutually exclusive.* [d] *Resting heart rate below the second percentile for age.* [e] *The same family member cannot be counted in A and B.* [f] *Definite LQTS is defined by an LQTS score ≥4. Scoring: ≤1 point, low probability of LQTS; 2–3 points, intermediate probability of LQTS; ≥4 points, high probability of LQTS.*

Source: Reprinted with permission from Hofman N, Wilde AAM, Kaab S, van Langen IM, Tanck MWT, Mannens MMAM, et al. Diagnostic criteria for congenital long-QT syndrome in the era of molecular genetics: do we need a scoring system? Eur Heart J 2007;28:575–80 and Schwartz PJ [20].

The International Long-QT Syndrome Registry was set up in 1979, and extensive clinical and genetic studies have been conducted since it began. Indeed, publications from the Registry have provided insight into risk mechanisms; genotype-phenotype associations; risk stratification by age, gender, and phenotype; and the importance of syncope as a cardiac event that frequently precedes aborted cardiac arrest or sudden cardiac death.

In 1991, Moss et al. [25] demonstrated the link between QTc and the risk of cardiac events. Zareba et al. subsequently examined the influence of genotype on the clinical course of the long-QT syndrome and found the risk of cardiac events to be higher among patients with a LQT1 mutation and those with a LQT2 mutation than those with a LQT3 mutation. However, the risk of lethal cardiac events was found to be highest in those with a LQT3 mutation [26].

Other studies that have investigated genotype-phenotype correlations and risk stratification in cLQTS include another study by Priori et al. [27]. This group developed a tool for gene-specific risk stratification with possible implications for disease management. They found the risk of cardiac events among those with a LQT1 mutation to be strongly dependent on QTc duration: males with a LQT1 mutation, and a QTc ≥500 ms, were at high risk for a first cardiac event in childhood, whereas for females with a mutation in the same gene and QTc in the same range, the risk of a first event remained unchanged over time [27]. Patients of either sex who had a mutation at the LQT1 locus and a QTc of <500 ms, had less than 30% risk of a first cardiac event before the age of 40 years. Female patients with a mutation at the LQT2 locus had a more severe prognosis irrespective of QTc duration. In patients with a LQT3 mutation prognosis was mainly influenced by sex, with male patients having a higher probability of becoming symptomatic by age 40 than female patients [27].

33.2.1.1 *Clinical Assessment and Management*

When a prolonged QTc is identified after a syncopal event in the absence of an acquired cause of prolonged QTc (see Section 33.3), the diagnosis of cLQTS should be considered, and family members can also be assessed. If there is a known pathogenic mutation in a patient originally diagnosed with cLQTS, then all first-degree relatives should be referred to a clinical geneticist and offered genetic testing for the familial mutation. Appropriate clinical follow-up of relatives should be arranged as required. Genetic testing of close family relatives is referred to as "cascade screening."

Sudden unexplained death in a young individual should be thoroughly investigated to determine the cause. If a structurally normal heart is identified at autopsy in a young person, conditions such as LQTS must be considered. Postmortem genetic testing is sometimes referred to as "molecular autopsy." LQTS represents a leading cause of autopsy-negative sudden death in the young [28].

Rarely, asymptomatic individuals with LQTS are identified when an ECG is obtained for other reasons (e.g., presurgery assessment) [10]. Gathering detailed information regarding family history in a suspected individual is essential because careful questioning may reveal a long-term pattern of similar episodes (syncope, sudden death), not only in first-degree relatives (mother, father, siblings, children) but also in more remote relatives in the family. Data on co-morbidities (such as congenital deafness) in evaluated individuals or family members should also be obtained [10]. It is important to distinguish, through careful history, acquired factors that result in QT prolongation from the inherited form of the condition [10]. Causes of abnormal prolongation of the QT interval include myocardial ischemia, cardiomyopathies, hypokalaemia, hypocalcemia, hypomagnesemia, autonomic influences, drug effects, and hypothermia (as discussed later in this chapter).

Beta-blockers are considered the first-line prophylactic therapy in LQT1 and LQT2. These drugs should be administered to all intermediate- or high-risk affected individuals and considered on an individual basis in low-risk cases [10]. Lifestyle and preventive measures such as the avoidance of potentially QT-prolonging drugs, adrenergic-type stimuli, and avoidance of alarm clocks should always be considered.

Some physician-investigators suggest beta-blocker therapy in all LQTS patients, even those at low risk (unless there is a contraindication), because of their mechanism of action and in view of the adrenergic-mediated triggers in this disorder, especially in those with LQT1 and LQT2. Implantable cardioverter-defibrillator (ICD) insertion can be highly effective in high-risk LQTS patients [10]. Surgical left-cervicothoracic sympathetic denervation, the surgical procedure introduced for treatment of LQTS before beta-blockers became available, can be considered in patients with recurrent syncope despite beta-blocker therapy and in patients who experience arrhythmia storms with an ICD [10]. Gene-specific LQTS therapies—sodium channel blockers, potassium channel activators, alpha-adrenergic receptor blockers, protein-kinase inhibitors, and atropine—have great potential and should also be considered.

TABLE 33.4 **Risk Factors for Drug-induced QT prolongation and Torsades de Pointes**

Female gender
Electrolyte disturbances • hypokalemia • hypomagnesemia • hypocalcaemia
Renal or liver dysfunction
Cardiac • myocardial infarction • congestive cardiac failure • bradycardia • recent cardioversion from atrial fibrillation
Congenital LQTS mutation carrier
LQTS gene polymorphisms
Drug-metabolizing enzyme polymorphisms

33.3 DRUG-INDUCED VARIATION IN QTc

Acquired long-QT syndrome has many different etiologies, including drugs, and several clinical conditions contribute, such as severe heart failure, cardiomyopathies, and left-ventricular hypertrophy [3]. For example, left-ventricular hypertrophy generates significant electrical remodeling, and its most characteristic feature is an extension of the duration of ventricular action potentials. This remodeling also increases the dispersion of repolarization and hence favors polymorphic ventricular tachycardia and sudden death [3]. See Table 33.4 for some of the risk factors for drug-induced LQTS. It has been demonstrated that each 10 ms increase in QTc interval contributes ~5% exponential increase in risk of cardiac events in patients with long-QT syndrome (LQTS) [29]. To further complicate matters, both cardiac and noncardiac drugs have been reported to prolong QT interval and induce arrhythmia in patients who have a QTc interval length within the reference range [8,30].

In 1964, Selzer and Wray [31] first described what would subsequently be termed *drug-induced LQTS* (diLQTS) and observed that quinidine could provoke QT prolongation and arrhythmias in otherwise healthy patients. The list of drugs causing QT prolongation includes those with no apparent cardiac action and those from different therapeutic families [32].

The Phenotype Standardization Project, sponsored by the International Serious Adverse Events Consortium (iSAEC) established a definition of the phenotype for drug-induced torsades de pointes (diTdP) to support research into pharmacogenomic markers of risk [33]. The proposed requirements for case classification are as follows:

- "The diLQTS event must have occurred during treatment with a known QT-prolonging drug, and/or a drug suspected of QT prolongation".
- "On withdrawal of the drug there is partial resolution of clinically significant QT prolongation (in the opinion of an arrhythmia specialist)".
- "Arrhythmia documentation must be reviewed by an appropriately experienced physician/electrophysiologist".

Three clinical subphenotypes for suspected diTdP were also defined to describe all possible presentations. Behr et al. [33] summarise them as follows:

- "Classical diTdP"– the pause-dependent onset of polymorphic VT (at least five beats) with associated QT prolongation documented with or without resuscitation and/or subsequent VF and cardiac arrest.
- "Probable diTdP" – polymorphic VT and/or VF documented but the typical onset of TdP has not been seen. QT prolongation must have been documented.
- "Possible diTdP" – ventricular arrhythmia has not been documented but a history of unexplained syncope has been elicited with severe QT prolongation on the ECG (QTc > 550 ms).

The absolute numbers of individuals exposed to risk of diTdP are significant, as up to 3% of all prescriptions worldwide include potentially proarrhythmic medications [5].

33.3.1 The Human Ether-à-go-go Gene Channel

The original name of the human ether-à-go-go channel (HERG) came from William Kaplan in the 1960s. Kaplan found that flies with mutations in this gene, when anesthetized with ether, showed movements in their legs similar to a popular dance called "à -go-go" [3]. The HERG is encoded in chromosome 7 and was first isolated in 1994 by Warmke and Ganetzky by screening a human hippocampal cDNA library with a mouse homolog of "ether- à -go-go" [3]. It is a member of the family of voltage-gated K^+ channels and has been shown to be responsible for the I_{Kr} [3]. HERG is composed of four identical alpha subunits (1159 amino acids) which co-assembles to form the full HERG channel. Each alpha subunit itself has six transmembrane domains, intracellular N- and C-termini, and a pore loop which links domains S5 and S6. Positive charges in the S4 domain act as voltage sensors. In the alpha subunits of the HERG channel, the S6 segment contains two aromatic amino acids (tyrosine-Tyr and phenylalanine-Phe). These are very important as they participate in the binding site of most of the drugs involved in QT prolongation.

In most K^+ channels, the selectivity filter is formed from the carbonyl atoms of a Gly-Tyr-Gly (GYG) sequence; however, in HERG channels the sequence is Gly-Phe-Gly (GFG) [3]. The central cavity is lined by the four S6 segments and is wide because these segments lack the bend found in other K^+ channels. This may explain why drugs of many different chemical structures can block HERG channels. The slow activation and rapid inactivation kinetics cause the HERG channel to pass significant current in the inward direction but little current in the outward direction (inward rectifier function). HERG mutations cause a reduction in outward K^+ current through multiple mechanisms—nonfunctional channel generation and mutations that affect intracellular trafficking of channel proteins [3].

The electrophysiological properties of the HERG potassium channel are important in many physiological processes and they may have different roles depending on where they are in the body [34]:

- Heart: they are the molecular basis for the cardiac muscles to rapidly produce I_{Kr} currents. Functional abnormality is related to long-QT syndrome (LQTS) and short QT syndrome (SQTS).
- Nervous system: they are involved in regulating spike-frequency adaptation and controlling resting potential. Abnormal expression is associated with the onset of schizophrenia.
- Cell proliferation and differentiation: they regulate apoptosis, and their abnormal expression is related to diseases such as cancer.
- Other: They are involved in regulating the secretory activity of pancreatic β cells and chromaffin cells. They also regulate the contractility of smooth-muscle cells in the portal vein, the jejunum, and the epididymal duct [34].

Most HERG-blocking drugs are compounds that cross the lipid bilayer as neutral molecules, equilibrate in the cytosol as positively charged molecules, and bind to the channel only in the open or inactivated state [3].

Tyr652 and Phe656 are the most important drug-binding sites. These high-affinity binding sites are located in the central cavity of the channel between the selectivity filter and the activation gate [3]. An inactivation-associated reorientation of Tyr652 and Phe656 during closure of the activation gate (formed by the crisscrossing of the S6 helical bundles) may cause drugs to be trapped inside the central cavity after inactivation. This process increases HERG channel drug sensitivity. On subsequent depolarization, because the drug remains bound, the channel is unable to conduct K^+ ions [3].

33.3.2 Repolarization Heterogeneity

There is intrinsic transmural heterogeneity in the density of the various ion channels, which has different effects on the action potential duration of the four cell types: subepicardial myocytes, midmyocardial M cells, subendocardial myocytes, and Purkinje cells [3]. Hence, drugs that only impair I_{Kr} preferentially prolong the action potential more in midmyocardium and Purkinje cells (which have lower densities of the other channels, such as I_{Ks} and I_{Na}). Thus, they increase heterogeneity of myocardium repolarization, which results in greater TdP risk. Amiodarone can produce marked prolongation of QTc without inducing TdP because it can block other ion channels, such as I_{Ca}, and I_{Na}, causing homogeneous prolongation of the action potential. More selective I_{Kr}-blocking drugs can increase heterogeneity in repolarization between the subepicardium and the endocardium [3]. This means that at a certain point the action potential in one location (e.g., the subepicardium) ends and action potential can be triggered by a new stimulus. However, in another location (e.g., the subendocardium) the cells are in early phase 3 and refractory, thus triggering reentry. It has been proposed that the last fragment of the T wave (T peak–T end) is a predictor of ventricular repolarization heterogeneity [3].

33.3.3 Mechanisms of Proarrhythmia in QT Prolongation

QT prolongation alone is not sufficient to result in TdP. Indeed, amiodarone usually results in significant QT prolongation, but TdP from amiodarone is rare [35,36]. Repolarization heterogeneity across cell types within the ventricular myocardium (see above) is being increasingly recognized as one explanation for the lack of direct link between QT prolongation and arrhythmias. The action potentials of the epicardial cells, midmyocardial (M) cells

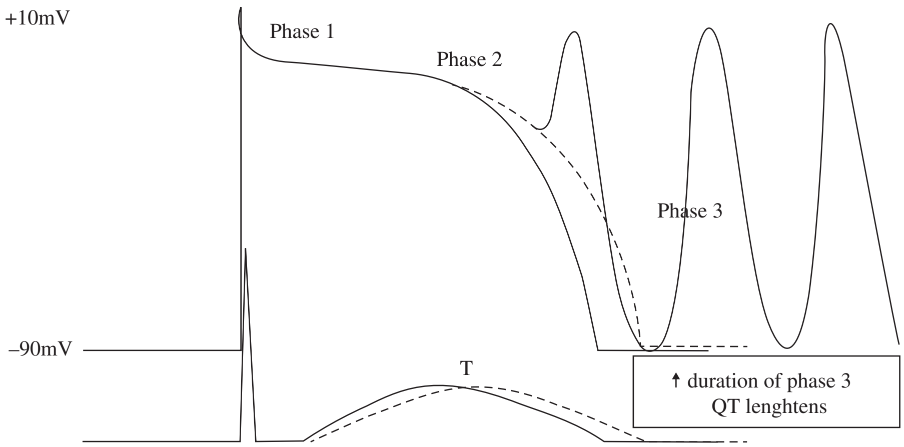


FIGURE 33.4 **Multiple EADs from progressively more negative transmembrane potential.** Prolonged repolarization can cause early after depolarizations (EADs) due to activation of inward depolarizing currents (L-type calcium channels or sodium-calcium exchange current), that appear as depolarizing oscillations in membrane voltage during phases 2 and 3 of the action potential. *Source: From Kallergis et al.* [121].

and endocardial cells differ in morphology and duration. These differences in turn are due to ion channel variations in the three cell types. To further complicate matters the action potential in each type responds differently to alterations in heart rate and to drugs [36].

This difference across the ventricular wall, otherwise known as transmural dispersion of repolarization, is physiologic; however, if exaggerated, it may increase TdP risk. Conversely, if transmural dispersion is reduced, as with the drug amiodarone which prolongs repolarization in all three cell types, this will lower the likelihood of TdP. The interval from the peak to the end of the T-wave on the surface ECG correlates with transmural dispersion of repolarization. Given the above information, this interval may signify a more specific ECG indicator of TdP risk in the presence of a long-QT interval [36].

Change in amplitude or polarity of the T-wave on alternating beats, otherwise known as T-wave alternans, is thought to result from alternation of the M-cell AP duration. It has been observed in LQTS as a TdP precursor because T-wave alternans leads to exaggeration of transmural repolarization dispersion during alternate beats, thus increasing the risk of TdP [37]. Another indicator of TdP risk is a change in QT duration on alternating beats or beat-to-beat QT variability. In an animal study, sotalol was used in a dose-dependent fashion to elicit TdP. Both low- and high-dose sotalol produced similar QT prolongation, however, "short-term" variability of the QT interval was increased only at the high dosage. This was accompanied by a high rate of TdP. In the same model, while amiodarone did prolong the QT interval, it did not increase short-term variability or result in TdP [36].

Action potential prolongation can lead to the development of early after depolarization (EAD) oscillations in the membrane potential during repolarization (see Figure 33.4). The danger inherent in a prolonged QT is that, if excessive, it carries a risk of sudden cardiac death (SCD) due to TdP. An EAD can result in an ectopic beat if it reaches a critical threshold in a large area of the myocardium. This ectopic beat is usually followed by a long pause, with a subsequent sinus beat showing marked QT prolongation. In the presence of exaggerated heterogeneity of action potential duration across the myocardium, the ectopic beat can induce reentrant excitation and TdP (see Figures 33.2 and 33.4). In contrast, TdP in its congenital form often follows a sudden adrenergic surge such as exercise or arousal [38].

EADs are thought to result from the reopening of either L-type calcium channels or sodium channels, and can result in ectopic beats if occurring in a large enough region of the heart. Triggered upstrokes from EADs are a likely initiating mechanism for TdP [36]. Proarrhythmia results indirectly from I_{Kr} blockade, which enables activation or reactivation of inward currents underlying EADs. This partly explains why verapamil, even though it is a potent I_{Kr} blocker, rarely causes TdP [6]. It has been shown to reduce EADs in LQTS, presumably by blocking the inward calcium current [39]. In a model of acquired TdP from a combined I_{Ks}–I_{Kr} blockade, verapamil has been found to shorten the QT interval, reduce transmural dispersion of refractoriness, and reduce EADs and TdP incidence [36]. A role for Ca/calmodulin-dependent protein kinase II (CaM kinase) in the development of TdP has also been demonstrated. Sparfloxacin was used to induce EADs in isolated rabbit Purkinje fibers, and was associated with severe prolongation of action potential duration and an increase in short-term repolarization instability. A CaM kinase block, caused by the calmodulin antagonist W-7, had no effect on action potential duration but did inhibit sparfloxacin-induced EADs in a concentration-dependent manner [36]. This decreased beat-to-beat repolarization instability.

33.3.4 Reduced Repolarization Reserve

The concept of "reduced repolarization reserve" suggests that multiple, often redundant, mechanisms maintain normal repolarization [6]. Minor alterations in function may not be obvious at baseline. A reduction in repolarizing current (e.g., as a result of a genetic variant that causes a small reduction in an important repolarizing ion channel), may not have any effect normally because other mechanisms maintain a normal QT interval [6]. The reduced reserve may, however, become more apparent when other factors influencing repolarization are added (e.g., slow heart rate, hypokalemia, or drug challenge). See Figure 33.5.

Initial findings leading to the concept of repolarization reserve began with the observation of quinidine syncope caused by TdP at a low plasma quinidine concentration [40]. There was no apparent explanation for this happening in a specific individual. In contrast, using sotalol therapy as an example, the risk of TdP was clearly increased as dose increased [4]. QT-interval prolongation reflects prolongation of at least some action potentials in the ventricular myocardium. Action potential prolongation must arise from an increase in inward current (e.g., sodium or calcium channels) or a decrease in outward current (e.g., through potassium channels) during the action potential [4]. Quinidine blocks both the transient outward and delayed rectifier components of the potassium current.

The two currents I_{Kr} and I_{Ks} have different activation, inactivation, and deactivation characteristics and different sensitivities to blocking drugs. In a later study, they were found to be the result of the expression of different genes—mutations that we now recognize as the common causes of congenital long-QT syndrome. Virtually all drugs that produce TdP are high-potency I_{Kr} blockers. Subsequent studies have identified molecular mechanisms underlying the susceptibility of this particular potassium channel to blocking by many structurally different compounds. There are only a few exceptions. For example, some studies have suggested that ibutilide prolongs QT interval and causes TdP by enhancing sodium entry into cells during the plateau [4].

The separation of I_{Kr} and I_{Ks} provided an explanation for unusual sensitivity to low concentrations of quinidine. The therapeutic range for quinidine is 2–5 $\mathrm{ug\,mL^{-1}}$ or ~6–20 $\mathrm{umol\,L^{-1}}$, and drug concentrations in the 5–10 $\mathrm{umol\,L^{-1}}$ range block multiple potassium currents as well as the inward sodium current [4,41–43]. However, when the effects of quinidine on the components of I_K were examined there was a difference: the concentration blocking 50% of I_{Kr} was 1 $\mathrm{umol\,L^{-1}}$, while I_{Ks} required 30–50 $\mathrm{umol\,L^{-1}}$ for a 50% block [4]. When extracellular potassium was lowered from 4 $\mathrm{mmol\,L^{-1}}$ to 1 $\mathrm{mmol\,L^{-1}}$, low drug concentrations of 0.4 $\mathrm{umol\,L^{-1}}$ (0.15 $\mathrm{ug\,mL^{-1}}$) were sufficient to produce a 50% I_{Kr} block. Quinidine's I_{Kr}-blocking actions are manifest at low concentrations; indeed, in some experimental settings increasing quinidine concentration blunts action potential prolongation. This might reflect superimposition of a sodium (or even calcium) channel block [44].

The principle that an I_{Kr}-blocking drug may not necessarily prolong action potentials if it also blocks inward current applies to other drugs such as the calcium channel blocker verapamil (as previously mentioned in Section 33.3.3). It is recognized as a potent I_{Kr} blocker but is not associated with QT prolongation or TdP [4]. Amiodarone also blocks I_{Kr} and prolongs QT interval [45,46]. However, amiodarone-induced TdP is unusual, presumably because it also blocks inward currents [35]. These findings support the idea that variability in the extent to which nonselective ion channel blockers like quinidine (or verapamil or amiodarone) prolong QT interval and generate TdP depends on their effects on other ion channels [4].

The data also suggested that the extent to which any I_{Kr} blocker prolongs QT may depend on the amplitude of other

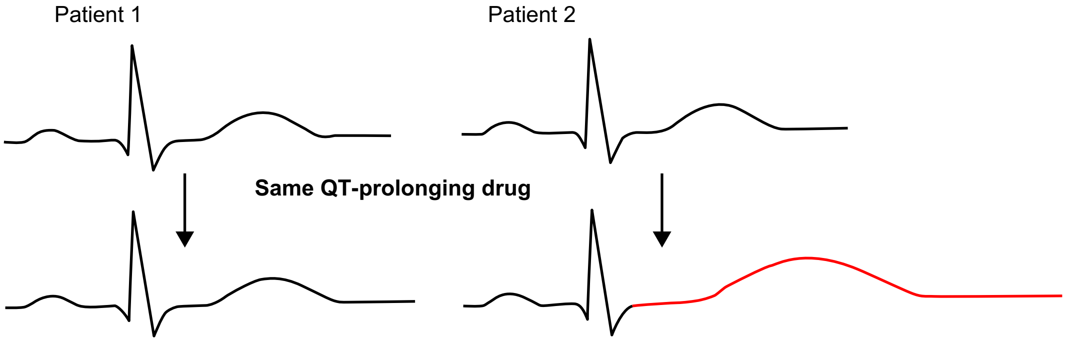


FIGURE 33.5 **Reduced repolarization reserve.** The concept of reduced repolarization reserve. Patient 1 is prescribed a QT-prolonging drug and has minimal or no prolongation of the QT interval. In contrast, patient 2, given the same QT-prolonging drug at the same dosage, develops marked prolongation of the QT interval. This has been termed "reduced repolarization reserve". Repolarization is a physiologically redundant process such that I_{Kr} block will not result in markedly prolonged repolarization; that is, the system displays some "reserve." However, otherwise subclinical lesions in other components of the repolarization system, such as reduction of I_{Ks} as a result of genetic factors or hypokalemia, may become apparent as marked QT prolongation when I_{Kr} is reduced. *Source: From Kannankeril et al.* [6] *with permission.*

ion currents, notably I_{Ks} [4]. Therefore, individuals with defective I_{Ks} might display normal QT interval because of a robust I_{Kr}, but are predicted to display marked QT prolongation when challenged with an I_{Kr}-blocking drug.

Roden summarizes well that the concept of repolarization reserve provides a "unifying framework for analysis of these risk factors and their clinical mechanisms" [4]. It is suggested that the normal heart includes multiple redundant mechanisms to accomplish normal repolarization, and that a lesion in just one of these mechanisms may not be enough to elicit a LQTS phenotype. However, the presence of subclinical lesions in the repolarization process means that superimposition of an I_{Kr}-blocking drug can produce marked action potential prolongation [4]. This can result in the pathophysiological cascade of engagement of inward current, increasing heterogeneities of action potential duration and EADs that culminate in TdP. This is a specific example of the idea that physiological redundancies in biological systems protect against extreme perturbations. Unusually high drug concentrations and toxicity are much more likely with drugs that are metabolized by a single (i.e., nonredundant) pathway than with drugs that are metabolized by multiple pathways [4].

Some patients may have very long QT intervals, and yet the vast majority of their heartbeats do not precipitate TdP. Therefore, (in keeping with the concept of repolarization reserve), in addition to genetic (or other) modulators of baseline QT interval and of the extent to which the QT interval is prolonged during drug exposure, it is also possible that genetic or environmental factors play a role in modulating the pro-arrhythmic pathophysiology that generates TdP when the QT interval is prolonged [4].

33.3.5 QTc Prolongation–Inducing Drugs—A Closer Look

The first recognition of drug-induced "acquired long-QT syndrome" came with the observation of syncope on quinidine therapy initiation (as mentioned in the previous section). This was recognized shortly after the introduction of the drug in the 1920s [4]. The first electrocardiographically documented case of acquired LQTS is thought to be the subject of a 1935 report in which hypomagnesemia was implicated [4]. Continuous electrocardiographic monitoring allowed for documentation of episodes of polymorphic VT in the 1960s, as reported for a series of patients on quinidine therapy [31]. Studies in the 1980s which revealed that quinidine blocks both the transient outward and the delayed rectifier components of potassium current [47,48]. and the identification of the delayed rectifier's two components, I_{Kr} and I_{Ks}, were key advancements in knowledge [49,50].

Most patients with diLQTS have had at least one of the risk factors listed in Table 33.4, in addition to drug exposure itself. Holter monitor studies of drug-induced TdP report an increase in the underlying sinus heart rate in the minutes prior to an event, suggesting that a pause in the setting of heightened sympathetic activation and long QT intervals may be especially arrhythmogenic [4].

The 2–3-fold female preponderance of QT prolongation [51] may be explained by the protective effect of testosterone. Electrocardiographical studies suggest that at puberty the QT interval shortens markedly in males and testosterone has been implicated as the major factor in the lowering of male TdP risk. Work in a prepubertal rabbit model reveals that an I_{Kr} blocker results in greater AP duration prolongation in males compared with females, and causes ventricular tachycardia in males only [6,36]. These findings may explain the higher risk of events in male children with congenital LQTS; however, other age- and gender-dependent factors may also contribute to diLQTS risk.

For the vast majority of drugs, risk increases with higher concentrations. Indeed, drug overdose and impaired drug elimination caused by genetic factors, disease, along with drug interactions, have been implicated as key mediators of TdP risk [4]. An interesting example is thioridazine, metabolized almost exclusively by cytochrome P450 CYP2D6 [52], which is functionally absent in 5–10% of Caucasians and African Americans and so results in higher plasma concentrations and elevated TdP risk. Here we discuss some more examples of drug-induced QT prolongation and TdP. See Table 33.5 for more information.

Antiarrhythmics

Antiarrhythmic agents are the leading cause of drug-induced TdP, and class IA agents (quinidine, procainamide, and disopyramide) block both Na^+ and K^+ channels and as previously discussed TdP can occur at either therapeutic or subtherapeutic doses [53]. Quinidine prolongs QT interval by an average of 10–15% within a week of therapy initiation and carries a 1.5% risk of inducing TdP [54]. TdP development on antiarrhythmic therapy is often precipitated by electrolyte disturbances, such as hypokalemia or hypomagnesemia. Potassium infusion has been found to reverse quinidine-induced QT prolongation [55]. Mortality increases threefold when quinidine is used in atrial fibrillation for rhythm control. Indeed, the incidence of quinidine-induced TdP ranges from 2.9-8.8% in patients with atrial flutter or atrial fibrillation and TdP tended to occur after conversion to sinus rhythm [55]. Due to its predominant Na^+-blocking effect, procainamide is less likely to induce TdP. However, in patients with renal dysfunction or in rapid acetylators, it can induce TdP through its active metabolite, N-acetyl procainamide (NAPA), which has a potent K^+-blocking effect [56]. Disopyramide, contraindicated in patients with heart failure and liver or renal dysfunction, has also been implicated in TdP [53].

Class III agents are potent I_{Kr} blockers and tend to prolong QT interval in a dose-dependent manner.[53], Ibutilide, dofetilide, and sotalol carry the highest risk for TdP, whereas amiodarone carries the lowest risk. Sotalol causes TdP in 2–4% of patients, with a higher risk seen in

TABLE 33.5 Drugs Reported to be Associated with QT Prolongation and/or TdP with Postulated Risk

Drug Group	High Risk (Frequent Reports of DITdP)	Moderate Risk (Reported to Cause diTdP)	Low Risk (diTdP likely Only with Overdose, Culprit Drug, or Concomitant Risk Factors)	Low Risk (QT Prolongation Only or Rarely with diTdP, Usually with Other Risk Factors)
Antiarrhythmics				
Class I	Quinidine,[a] procainamide,[a] Disopyramide, dihydroquinidine[a]	Flecainide, propafenone, pirmenol, [a] cibenzoline, [a] ajmaline[a]		
Class III	Sotalol, d-sotalol,[a] dofetilide,[a] azimilide,[a] Ibutilide,[a] sematilide,[a] ersentilide,[a] almokalant,[a] nifekalant,[a] terikalant[a]	Amiodarone	Dronedarone	
Antianginals and vasodilators	Prenylamine,[a] terodiline,[a] lidoflazine,[a] bepridil[a]			Vardenafil, ranolazine
Antihypertensives		Indapamide		Nicardipine, isradipine, moexipril/hydrochlorothiazide
Anti-histamines	Terfenadine,[a] astemizole[a]	Ebastine[a]	Diphenhydramine	Mizolastine, bilastine
Serotonin agonists and antagonists	Cisapride,[a] ketanserin[a]	Dolasetron		Ondansetron, granisetron
Antimicrobials				
Macrolides	Erythromycin, spiramycin	Azithromycin, clarithromycin	Roxithromycin[a]	Telithromycin
		Gatifloxacin,[a] levofloxacin, Sparfloxacin,[a] moxifloxacin, grepafloxacin[a]		
Fluoroquinolones			Ciprofloxacin	Gemifloxacin,[a] ofloxacin, gatifloxacin
Antifungals			Ketoconazole, fluconazole, itraconazole, voriconazole	Posaconazole
Antimalarials	Halofantrine[a]	Chloroquine	Quinine	artemether with lumefantrine,
Others	Pentamidine		Amantadine, cotrimoxazole,	Foscarnet, atazanavir,
			Trimethoprim, ritonavir	Saquinavir, telaprevir
Psychiatric				
Tricyclic antidepressants			Amitriptyline, nortriptyline, desipramine,[a] clomipramine, imipramine, trimipramine, doxepin, protriptyline[a]	

TABLE 33.5 Continued

Drug Group	High Risk (Frequent Reports of DITdP)	Moderate Risk (Reported to Cause diTdP)	Low Risk (diTdP likely Only with Overdose, Culprit Drug, or Concomitant Risk Factors)	Low Risk (QT Prolongation Only or Rarely with diTdP, Usually with Other Risk Factors)
Serotonin reuptake inhibitors		Citalopram	Fluoxetine, paroxetine, sertraline, zimeldine,[a] amoxapine[a]	Venlafaxine, escitalopram
Antipsychotics	Thioridazine,[a] chlorpromazine, haloperidol, droperidol[a]	Pimozide, mesoridazine, sertindole	Amisulpride, ziprasidone, risperidone	Trifluoperazine, prochlorperazine, Paliperidone, quetiapine, flepentixol, fluphenazine, zuclopenthixol
Others			Maprotiline,trazodone	Lithium
Anticancer	Arsenic trioxide	Tacrolimus, geldanamycin,[a] vandetanib		Nilotinib, tamoxifen, sunitib, lapatinib, vinflunine, erubilin, pazopanib, fingolimob, capecitabine
Others	Probucol, domperidone, cisapride, methadone	Vasopressin, tizanidine, alfuzosin, amantadine, levomethadyl,[a] lofexidine	Organophosphates,[a] galantamine, solifenacin, clobutinol,[a] chloral hydrate, sevoflurane	Felbamate,[a] fosphenytoin, degarelix, perflutren, octreotide, oxytocin, galantamine, retigabine, tizanidine, tolterodine, alfuzosin, isolfurane

[a] *Unlicensed, withdrawn, or suspended in the United Kingdom.*
Source: Behr ER and Roden D [5], with permission.

women [57]. The incidence of TdP with Dofetilide is 2.1%, with increased risk in renal failure patients, and intravenous ibutilide carries a TdP risk of 1–3% [58], with a higher incidence in patients with structural heart disease, heart failure, and electrolyte disturbance. The incidence of TdP with amiodarone is rare despite its QT-prolonging effect [53].

Antihistamines

Only a few nonsedating antihistamines have been implicated in significant arrhythmogenicity. Due to their potent I_{Kr}-blocking effect even at low doses, terfenadine and astemizole have been withdrawn from the market because of their high TdP risk. Most antihistamines are metabolized by the cytochrome P450 enzyme CYP3A4. Patients with liver dysfunction or who are co-administered drugs or food (such as grapefruit) that inhibit CYP3A4 can show higher drug levels. The newer nonsedating antihistamines such as desloratadine, cetirizine, levocetirizine, and fexofenadine seem to be safer than previous versions [53,59].

Antipsychotics and antidepressants

Antipsychotic medications prolong QT interval in a dose-dependent manner [53]. Haloperidol (a butyrophenone) is used to treat schizophrenia and severe agitation and is known to be a potent blocker of the I_{Kr} channel. This effect is amplified in the presence of risk factors (see Table 33.4). Haloperidol has been reported as prolonging the QT interval by up to 15–30 ms [53] and in 2007, the US Food and Drug Administration (FDA) released an alert recommending ECG monitoring with intravenous haloperidol use. Droperidol has an effect similar to that of haloperidol and, in patients with risk factors, clinicians must pay careful attention and consider whether ECG monitoring may also be appropriate with its use.

Phenothiazines such as chlorpromazine have a combination of both antipsychotic and antiemetic effects [53]. Indeed, both chlorpromazine and thioridazine have been implicated in QT prolongation and TdP due to their K^+-blocking effect. Atypical antipsychotics can cause dose-dependent prolongation of the QT interval associated with variations in their potency. QT prolongation is highest with ziprasidone and lowest with olanzapine [60]. In 1998, sertindole was withdrawn from the market due to the risk of TdP and sudden death. Zimelidine and citalopram have also been related to TdP in high doses [53].

Tricyclic antidepressants (TCAs) are more commonly associated with QTc-interval prolongation than selective serotonin-reuptake inhibitors (SSRIs) [61]. The mechanism by which TCAs prolong the QTc is blockade of the Na^+ channel. The effect is exaggerated when there is co-administration of a potassium channel–blocking agent [53]. Amitriptyline, desipramine, and imipramine have all been implicated in TdP, and toxic doses of TCA, as seen in TCA overdose, can result in various ECG changes such as QRS complexes widening, QT prolongation, and TdP [53]. All such cases should have ECG monitoring when being assessed in the acute phase. SSRIs prolong the QT interval by inhibiting the I_{Kr} channel, and citalopram and escitalopram have been associated with QT prolongation [53].

Antibiotics and antimalarials

As Nachimuthu et al. state [53], fluoroquinolones have a variable effect on the QTc interval and bring with them a very low incidence of TdP. Grepafloxacin and sparfloxacin delay repolarization more profoundly than gatifloxacin, levofloxacin, and moxifloxacin, with ciprofloxacin and ofloxacin having the least effect on the I_{Kr} channel [53]. Indeed, both grepafloxacin and sparfloxacin were discontinued in the preliminary drug development. The rest of the fluoroquinolones are relatively safe, but caution must be exercised if there are any underlying risk factors or if they are co-administered with other QT-prolonging drugs [62].

Macrolides, (e.g. clarithromycin and erythromycin), have also been associated with QT prolongation and TdP [61]. In animal studies, erythromycin was found to have similar effects to those of class III antiarrhythmic agents, with prolongation of QT interval and EAD induction (see Figure 33.4.). Both erythromycin and clarithromycin are CYP3A4 inhibitors which can unsurprisingly lead to significant toxicity if co-administered with another CYP3A4 inhibitor or with drugs metabolized by CYP3A4 [53]. TdP has also been reported with the use of azithromycin, although the drug is considered to be safe [63].

The true incidence of QT prolongation and TdP with antimalarial agents is largely unknown. This is because of underreporting on a worldwide level [53]. Quinine, which is an optical isomer of quinidine, has a weak effect on the QT interval and rarely has been associated with TdP. However, chloroquine and halofantrine can prolong the QT interval and so have been linked to TdP. Halofantrine the most potent, has been found to have repolarization properties similar to those of quinidine and Class III antiarrhythmic agents [53,64].

Others

Pentamidine is an anti-protozoal drug that is used for *Pneumocystis carinii* pneumonia treatment [53]. It has structural similarity to procainamide and while its electrophysiological properties are not known, intravenous use can lead to TdP. Inhalation is considered safe. The proarrhythmic risk of IV use is related to idiosyncratic reaction rather than the effect of dose [65].

The systemic azole group of antifungal agents have both pharmacodynamic and pharmacokinetic characteristics that may trigger TdP [53]. Attention should be paid when using these agents, especially in patients who have underlying risk factors for QT prolongation (see Table 33.4).

Cisapride, the gastrointestinal promotility agent used for delayed gastric emptying and gastroesophageal reflux disease [53] is also structurally similar to procainamide and has demonstrated both I_{Kr-} and I_{Ks-} blocking effects. Cisapride use has been related to the greatest number of TdP cases next to antiarrhythmic agents [66] which ultimately led to its withdrawal from US market in July 2000.

33.3.6 Pharmacokinetic and Pharmacodynamic Factors

Pharmacokinetic and pharmacodynamic interactions might increase the risk of TdP of torsadogenic drugs. Pharmacokinetic interactions occur mainly because of the inhibition or induction of cytochrome P450 enzymes [3]. Pharmacokinetic interactions may occur if one drug reduces the clearance of the other or if both drugs compete for the same hepatic enzyme. Co-administering a drug with cytochrome CYP3A4 inhibitors such as -azoles, -mycins, or grapefruit juice will increase its level. Haloperidol and thioridazine share the same enzyme, CYP2D6, for clearance, resulting in increased levels of both [53,66,67].

Terfenadine undergoes near-complete presystemic metabolism by CYP3A4 (and the closely related isoform CYP35) in the intestine and liver. As a result, plasma concentrations of the drug are low, and most of its antihistamine actions are mediated by generation of an active metabolite, fexofenadine (now on the market as Allegra) [4]; unlike with CYP2D6, patients with no CYP3A activity on a genetic basis have not been recognized. Terfenadine itself is a potent I_{Kr} blocker and, with inhibition of CYP3A, causes markedly elevated plasma concentrations and TdP risk [4].

TdP-related pharmacodynamic interactions are easier to detect because they occur when two drugs act on the same or a distinct target and induce the same adverse drug effect, such as QT prolongation (e.g., antipsychotics and tricyclic antidepressants) [3]. Other mechanisms include the combination of a QT-prolonging drug with another drug that influences a variable that increases QT prolongation risk. Examples of such drugs include those that produce hypokalemia, such as beta 2 agonists, diuretics, amphotericin B, and insulin.

33.3.7 Prevention, Monitoring and Treatment of Drug-Induced LQTS

QT-prolonging drugs should be avoided in patients with preexisting heart disease, with a history of ventricular arrhythmias or with metabolic abnormalities such as hypokalemia (see Table 33.4). Hospitalized patients are at a higher risk for developing TdP than outpatients with the same QT-prolonging drugs [53]. This is because hospitalized patients are often elderly with underlying heart disease, who may also have renal or hepatic dysfunction, electrolyte abnormalities, and bradycardia, to whom drugs need to be administered intravenously and rapidly [68]. Co-morbidities act as, or enhance, risk factors for TdP. Concomitant administration of drugs that inhibit the cytochrome P450, especially imidazole antifungals, and macrolide antibiotics or those that can prolong the QT interval or cause electrolyte disturbance, should be avoided [53]. Surveillance ECGs are recommended before and after initiation of QT-prolonging drugs. Routine electrolyte monitoring, especially for potassium, is also recommended in those who are on QT-prolonging drugs and diuretics. For example, methadone has been associated with TdP in women if the dose is >100 mg/day [69]. Methadone guidelines recommend a pretreatment ECG for QTc-interval screening and a follow-up ECG within 30 days and then annually [53,70].

ECG monitoring of the QT interval in the hospital setting is indicated for the following reasons: (1) initiation of a drug known to cause TdP; (2) overdose from potentially proarrhythmic agents; (3) new-onset bradyarrhythmias; and (4) severe hypokalemia or hypomagnesemia [53]. If drug-induced TdP has occurred, a careful review of the patient's personal/family history must be obtained in order to identify the possibility of congenital LQTS. If a history of unexplained syncope or premature sudden death is discovered, 12-lead ECG is recommended for all first-degree relatives and genetic testing for congenital LQTS should be considered in the family.

In cases of TdP which have degenerated into ventricular fibrillation, immediate direct current cardioversion should be performed [53]. In those with TdP who are otherwise stable, the following treatments should be considered: intravenous magnesium, intravenous potassium, overdrive transvenous pacing. A detailed review of the patient's medication list should be carried out. Permanent pacemakers should be inserted if the patient has chronic bradycardia e.g. AV block or sick sinus syndrome, however, implantable cardioverter defibrillators (ICDs) should also be considered, and the case should therefore be discussed with an expert in the field.

33.3.8 Detecting Drug-Induced Effects on Repolarization

There are many models to detect if a drug can potentially prolong QT [3]. Several examples of *in vitro* models are described to evaluate their effects on repolarization. Some examples are heterologous expression systems (in studying ion currents), disaggregated cells (also in studying ion currents), isolated tissue (in screening large numbers of compounds), and isolated intact heart (also in screening large numbers of compounds) [3].

In vivo models include multi-lead ECG in conscious or anesthetized guinea pigs, rabbits, dogs or pigs. The QT should be measured from at least three successive beats. However, there is no single, standardized, uniform method for preclinical screening of drug effects on cardiac repolarization [3].

33.3.9 Congenital LQTS Genes in Drug-Induced LQTS

The accumulation of case series of diTdP has identified small numbers of subjects in whom the diagnosis of congenital LQTS was made only after culprit drug exposure and arrhythmia [71]. This is consistent with the concept that some patients with congenital LQTS have little QT prolongation at baseline but may be more susceptible to a markedly prolonged QT interval on exposure to a QT-prolonging drug [5,23]. Around 10% of subjects with diLQTS have rare variants in the cLQTS disease genes and are thought to represent incompletely penetrant cLQTS. Studies on this focused on the more common disease genes in cLQTS [72,73], such as *KCNQ1* and *KCNH2* and not on the rarer, more recently identified genes. In some recent studies, individuals with subclinical cLQTS or genetic polymorphisms have been found to be more predisposed to drug-induced or hypokalemia-induced arrhythmia [74]. In cLQTS, *KCNQ1* mutations are most common, and, as we know (see Table 33.2) this gene encodes I_{Ks}. Loss of I_{Ks} function may not be apparent until superimposition of I_{Kr} block. Genetic polymorphisms that have been found in the sodium channel produce subtle changes in sodium channel activation until the I_{Kr} blocker is superimposed, at which point exaggerated action potential prolongation and EADs are manifest [4,74].

33.3.10 *KCNH2* Polymorphisms and Disease Susceptibility

Since the early finding in 1995 that functional deficits of HERG potassium channels are related to cLQTS [34], over 293 LQTS-related *KCNH2* mutations have been reported in the literature, including a new frameshift mutation recently discovered by Mihic et al. (P1086fs+32X, 3256InsG) [75]. The Arg744Pro mutation found by Aidery et al. [76] and the Gly816Val mutation reported by Krishnan et al. [77] are both associated with cLQTS. Copy number variation (CNV) in *KCNH2* may also be involved in LQTS. For example, Barc et al. [78] recently reported that *KCNQ1* and *KCNH2* CNVs might account for 3% of LQTS occurrence.

The clinical symptoms of aLQTS are similar to those of cLQTS. aLQTS is caused by environmental factors, including cardiomyopathy, abnormal metabolism, and, most commonly, drugs. It manifests as prolonged QT interval, sometimes accompanied by TdP and ventricular fibrillation [34]. Susceptibility to aLQTS significantly correlates with gene polymorphisms. For example, the mutation Pro347Ser is related to cisapride- and clarithromycin-induced TdP; Arg784Trp, with amiodarone-induced TdP.

33.4 PUBLIC HEALTH IMPACT

Long-QT syndrome is a significant public health problem, especially given the large number of drugs that can prolong the QT interval and its risk of perhaps a fatal outcome. Vast numbers of patients worldwide are exposed to potentially QT-prolonging agents every day, and while the actual number of those suffering adverse outcomes is small, the severity of outcome in those affected keep this issue a priority. It is for this reason that the QT interval is thoroughly evaluated as part of the drug safety and development process (see Section 33.5).

TdP arising during antiarrhythmic drug therapy is an important consideration when balancing risk and benefit in individual patients and in large numbers of subjects exposed to these drugs in clinical trials and practice. The risk of TdP versus the potential benefit of anti-arrhythmic therapy was the centerpiece of the FDA advisory panel debates over ibutilide, dofetilide, and sotalol [4]. D-sotalol, which is the dextrorotatory isomer of sotalol, was developed as a non-beta-blocking "pure" class III anti-arrhythmic. However, it was abandoned when the SWORD (Survival With Oral D-Sotalol) study demonstrated increased mortality with the drug versus placebo [4]. Drug-induced TdP was one possibility. Mibefradil, a drug with L- and T-type calcium channel–blocking properties, increased mortality in patients with heart failure, compared with placebo [79]; the fact that there was an increased risk among female patients receiving diuretics, raises the possibility that TdP may have been implicated.

The possibility that TdP also occurs during drug therapy not destined for cardiovascular indications was recognized early in the 1970s, with antipsychotic agents [80], and in the 1990s, with the very widely used antihistamine terfenadine [81,82]. By the time the potential for TdP was recognized, terfenadine had been used in tens of millions of patients and had generated such a good safety record that over-the-counter status was being considered [4]. It was the recognition of a single case of TdP, associated with the concomitant use of terfenadine and ketoconazole, that prompted a review of the FDA voluntary reporting postmarketing database, which identified several dozen more cases. A common feature of these cases was co-administration of drugs known to inhibit certain pathways of drug metabolism (e.g. CYP3A4) [4].

Initial attempts to deal with the problem focused on educational programs to prevent co-prescription of terfenadine with known inhibitors such as erythromycin and ketoconazole, but these programs were not sufficiently effective and the drug was withdrawn [4]. While terfenadine, and other drugs since, have received considerable attention in the press, the incidence of documented TdP associated with these drugs remains much lower than that associated with anti-arrhythmic drugs. However, as in the case of terfenadine, given that there were effective alternative therapies and that the indication for therapy was not in itself life-threatening, even a very small risk of TdP was considered to be unacceptable.

The absolute numbers of individuals exposed to diTdP risk are significant as up to 3% of all prescriptions worldwide include potentially pro-arrhythmic medications [83]. Therefore, diTdP has become a very important public health and regulatory issue and has led to the withdrawal of a number of widely prescribed drugs from the market, such as, cisapride and thioridazine, along with terfenadine. It has also led to restricted use of others (e.g., citalopram, droperidol, haloperidol, sertindole) [5,84].

Some pro-arrhythmic drugs that are essential to medical practice have not been withdrawn, such as macrolide antibiotics in penicillin-sensitive individuals and halofantrine in drug-resistant malaria [5]. Several neuroleptic drugs, vital in the treatment of psychotic illnesses also remain in use.

With the recognition of terfenadine-induced TdP came the increasing scrutiny of other drugs and it is now routine for all new drug entities to be screened for their ability to block I_{Kr} *in vitro* and to prolong QT intervals in early human trials. Although this has imposed a new hurdle in drug development, screening studies will reduce the likelihood of any new drug unexpectedly causing TdP by blocking I_{Kr} [5,8,83–85].

Clinicians are often faced with decisions regarding alternative medications for individuals with a prolonged QT interval. However, cardiovascular screening prior to prescription of potentially QT-prolonging drugs seems unlikely to be cost effective or widely adopted. Still, clinicians should be conscious of cardiac-related aspects of the patient history and should consider any potential increased risks prior to commencing such therapies.

Specific groups of patients, particularly those with risk factors as mentioned in Table 33.4, should be observed prior to therapy with potentially QT-prolonging agents, ideally in a medical setting. Also, patients whose family history includes sudden cardiac death or those on complex medical regimes that may give rise to issues of drug elimination should be treated with caution. In reality, this is often not possible, and when one considers the great number of patients fulfilling such criteria who are treated every day, both in hospital and in the community, versus the number who will suffer an episode of TdP, it may be difficult to justify such extra measures on a cost-effective basis.

Nevertheless, if therapy with a potentially QT-prolonging agent is commenced, the patient should be warned to promptly report any unusual symptoms, particularly any cardiac-related symptoms such as palpitations and syncope [30]. It is also important to remember that other intercurrent symptoms may alter electrolytes and predispose to cardiac events—for example, gastroenteritis, which may lower serum potassium levels. An ECG during treatment may detect the asymptomatic prolongation of the QTc to >500 ms or an increase in the QTc of >50 ms on treatment. In this, education programs about co-prescribing have not been sufficiently effective [4]. It is hoped that increased awareness among clinicians (especially family doctors, who are generally the first-line prescribers) and patients themselves may result in greater exposure to highly informative sources such as www.qtdrugs.org. A higher profile for cardiac-related and sudden death charities (and the information they provide) on social networking sites may prove to be beneficial in the long term. Improved clinician education and awareness of the proarrhythmic potential of noncardiac drugs alone is probably not sufficient in preventing adverse outcomes, and computer-assisted strategies can offer another level of safeguarding. Electronic prescribing allows for sophisticated systems with automatic alerts and this will add an extra level of defense against prescribing medications which may be proarrhythmic in combination, or in conjunction with a patient's current medical history e.g. temporary electrolyte imbalance whilst a hospital inpatient. Systems which can combine a patient's past medical history with their current "real time" medical experience e.g. hypokalaemia or heart failure would be most beneficial.

A major concern is always the fact that rare and poorly understood side effects can occur with many highly effective drugs, and that the withdrawal of these medications from the market altogether would probably harm more patients than it would help. The very real risk of serious adverse events in the minority must be weighed against the clear therapeutic benefits in the vast majority.

33.5 QTc IN NEW DRUG DISCOVERY AND DRUG SAFETY

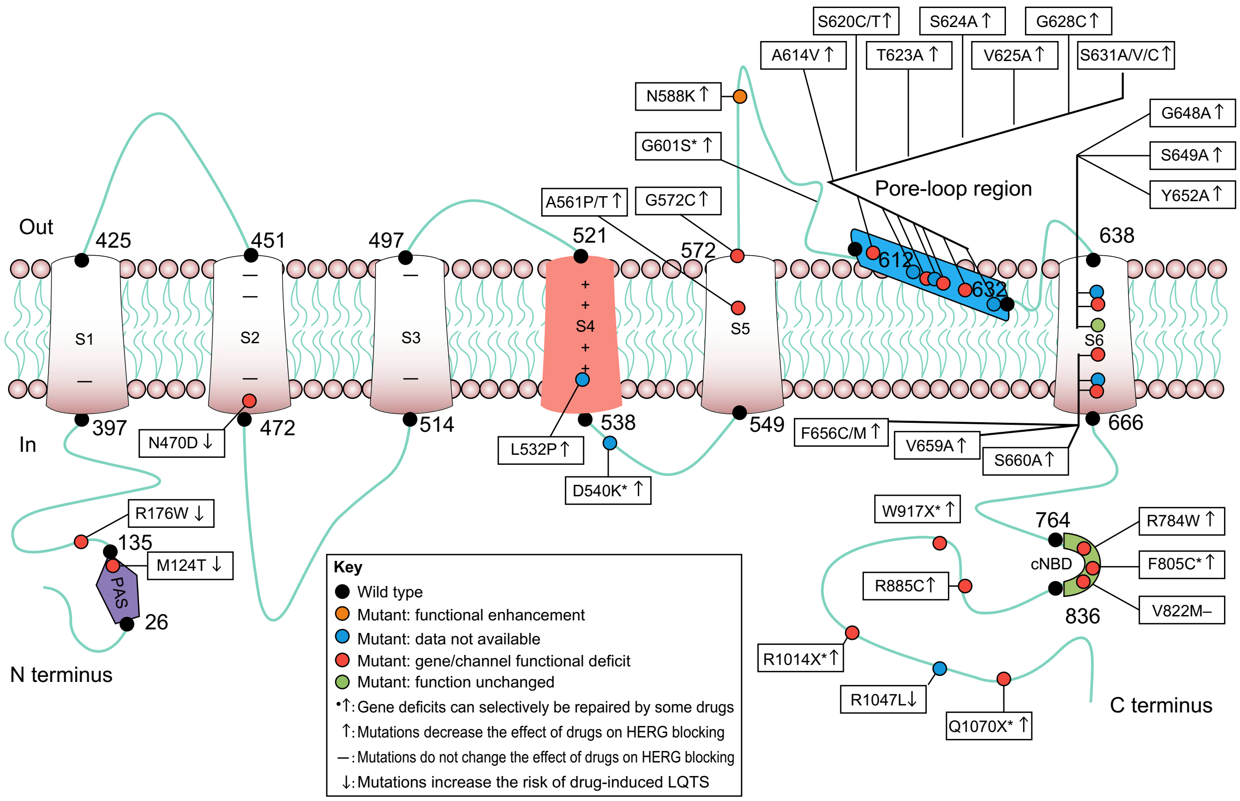
As Nachimuthu et al. summarize well, it has now become common practice for pharmaceutical and biotechnology companies to screen compounds for HERG channel activity early in preclinical safety assessment [53]. Indeed, in Europe and the United States all new molecules must undergo evaluation in both pre-clinical and clinical settings to assess their effects on repolarization. An example of this in practice is the "Thorough QT/QTc (TQT) study", intended to confidently identify drugs that may cause QT prolongation [53]. This study has evolved as a key component of all clinical development programs for new molecular entities; it employs a double-blinded, randomized design with a placebo and a positive control arm that is strictly powered to exclude an effect on the QTc interval exceeding 10 ms. It aims to determine whether the new chemical entity prolongs the QTc interval, and if so, whether the effect is dependent on the dose or the concentration. It is conducted using healthy volunteers once the tolerability and the pharmacokinetics of the drug have been established. During a TQT study, RR interval, heart rate, QT interval corrected using both Bazett's and Fridericia's correction are collected. ECG interval measurement methods are either fully manual or use some degree of manual adjudication (e.g. computer-based algorithmic measurements are reviewed and corrected as needed) [53].

The role of the positive control is to demonstrate the study's ability (i.e., assay sensitivity) to detect a small effect on the QT interval—in this case, QT prolongation of slightly more than 5 ms. If the study can detect such a small QT prolongation by the control, then a finding of a lesser QT effect for the test drug should indicate that the drug does not significantly prolong the QT interval. Moxifloxacin, a fluoroquinoline antibiotic with a mild QT-prolonging effect is used in most of these studies [53]. Moxifloxacin 400 mg has been associated with an observed 7.5–12.5 ms increase in the mean placebo- and baseline-corrected QTc interval, which supports its appropriateness a positive control in TQT studies [86].

A TQT study usually results in three clinical scenarios [53]:

1. The studied drug prolongs the mean QTc interval by ≤10 ms and therefore does not appear to cause TdP.
2. The studied drug prolongs the mean QTc interval by >10 ms but ≤20 ms and has an "uncertain" risk of TdP.
3. The studied drug prolongs the mean QT/QTc interval by >20 ms and may have a "high" risk for causing clinically important arrhythmic events.

To fully evaluate the risk of QT prolongation, it is important to include patients in these analyses with additional TdP risk factors such as those with electrolyte abnormalities (e.g., hypokalemia), congestive heart failure, impaired drug-metabolizing capacity or clearance (e.g., renal or


S620C/T↑
S624A↑
G628C↑
A614V↑
T623A↑
V625A↑
S631A/V/C↑
N588K↑
G601S*↑
G648A↑
S649A↑
Y652A↑
Pore-loop region
A561P/T↑
G572C↑
Out
425
451
497
521
572
612
632
638
S1
S2
S3
S4
S5
S6
In
397
472
514
538
549
666
N470D↓
L532P↑
D540K*↑
F656C/M↑
V659A↑
S660A↑
R176W↓
135
M124T↓
PAS
26
N terminus
W917X*↑
764
R784W↑
cNBD
F805C*↑
836
V822M–
R885C↑
R1014X*↑
R1047L↓
Q1070X*↑
C terminus
Key
● Wild type
● Mutant: functional enhancement
● Mutant: data not available
● Mutant: gene/channel functional deficit
● Mutant: function unchanged
•↑: Gene deficits can selectively be repaired by some drugs
↑: Mutations decrease the effect of drugs on HERG blocking
—: Mutations do not change the effect of drugs on HERG blocking
↓: Mutations increase the risk of drug-induced LQTS

TRENDS in Molecular Medicine

FIGURE 33.6 **HERG subunit.** Shown here is a single HERG subunit containing six α-helical transmembrane domains (S1–S6), a pore-forming loop (between S5 and S6), a C-terminal cyclic nucleotide-binding domain (cNBD), and a N-terminal Per-Arnt-Sim (PAS) domain. There are at least 30 *KCNH2* SNPs shown associated with HERG potassium channel sensitivity to about 70 drugs; polymorphisms leading to changes in the channel properties are highlighted. *Source: From He et al.* [34] *with permission.*

hepatic impairment, drug interactions), female patients, and patients aged <16 years and >65 years.

Recent studies have indicated that fully or partially automated methods of QT interval measurements using computer algorithms, compared with labor-intensive manual methods, produce similar results when tested on drugs with a known QT-prolonging effect [87]. It is expected that use of automated QT algorithms in the evaluation of new medications will not only replace manual QT measurements but will also improve detection of subtle T-wave changes [53,88].

33.6 PHARMACOGENOMIC STUDIES

33.6.1 HERG Potassium Channels, *KCNH2* Polymorphisms, and Drug-Induced LQTS

Figure 33.6, by He et al. [34] shows a single HERG subunit, with 30 *KCNH2* single-nucleotide polymorphisms (SNPs) that are associated with HERG potassium channel sensitivity to about 70 drugs [34]. These drugs include antiarrhythmic agents, antidepressants, and antibiotics. The following Sections 33.6.1–33.6.3 describe the relationship between *KCNH2* polymorphisms and drug-induced QT-interval abnormalities according to the literature reported and summarized by He et al. [34].

Current studies on HERG potassium channel polymorphisms and drug-induced LQTS began with finding drugs related to QT interval prolongation. *In vitro* patch-clamp techniques were then used to determine the effect of these drugs on I_{Kr} currents in recombinant cells which had different HERG potassium channel genotypes. In this way, the relationship between *KCNH2* polymorphisms and drug-induced QT abnormalities could be evaluated. However, one must accept that drug-induced malignant arrhythmia is rare and the numbers of LQTS patients are relatively small [34]. Nevertheless, with new technological developments such as induced pluripotent stem cells (iPSCs) [34,89], cell models that closely mimic the electrophysiological properties of cardiac-muscle cells can be employed to evaluate the effects of HERG potassium channels more objectively.

At present, *in vitro* studies have demonstrated that the majority of LQTS-related drugs block HERG potassium channels, and HERG potassium channel polymorphisms affect the sensitivities of these drugs [34]. In particular, the Tyr652Ala and Phe656Ala mutations in the S6 pore region of HERG potassium channels are the key to drug sensitivity [34]; however, these results have not as yet been clinically proven. Interestingly, natural plant monomers, such as berberine, nicotine and papaverine can lead to prolonged QT intervals or action potential duration (APD), and *KCNH2* polymorphisms are related to the sensitivity of these traditional Chinese medicines [34].

33.6.2 Functional Deficits Caused by HERG Channel Mutations

Congenital LQTS induced by *KCNH2* polymorphisms is mainly caused by functional deficits or transfer deficits of HERG potassium channels [34]. Mutations at the C terminus of HERG potassium channels can cause misfolding, transfer deficit, and abnormal positioning and thus disturb the ion distribution density needed for these channels to function normally. This can lead to LQTS (e.g., the Asn470Asp mutation) [34]. However, ion channel functional deficits or transfer deficits caused by these mutations can be repaired by HERG blockers or by low temperature. For example, Harley et al. [90] found that the transfer deficits caused by nine mutations in the PAS region of HERG potassium channels could be rescued by low temperature [34].

Delisle et al. [91] found that thapsigargin can selectively rescue the transfer deficit of HERG potassium channels caused by Gly601Ser and Phe805Cys mutations. It was also found that the HERG blocker E4031 can repair deficits caused by Asn470Asp and Gly601Ser [34]. Yao et al. showed that current density decreases caused by Arg1014X and Trp927X mutations could be partially repaired by gentamicin [34,92]. Rajamani et al. found that, although rescuing the transfer deficits caused by Asn470Asp and Gly601Ser mutations does not depend on the HERG-blocking effect of terfenadine and fexofenadine, the transfer deficit caused by Val822Met mutation could not be repaired [34,93].

Gong et al. [34,94] showed that morpholino oligonucleotides can counteract downregulation of the mRNA level induced by the HERG Gln1070X mutation and increase HERG protein expression. The researchers indicated that investigating *KCNH2* polymorphisms and QT syndromes will be of clinical value in developing individualized treatments.

33.6.3 Distribution of Drug Sensitivity–Related *KCNH2* Polymorphisms

Studies on the frequency distribution of *KCNH2* polymorphisms in human populations have been rare. Also, there are only a few studies on the polymorphisms related to the drugs that act on HERG potassium channels. Mutations in HERG that correlate with drug sensitivity are many, but there have been only a few studies on the frequency distribution of *KCNH2* polymorphisms in human populations [34].

The reported drug sensitivity–related *KCNH2* polymorphisms show substantial racial differences, which suggests that studies on genes related to HERG drug sensitivity in human populations are important [34].

HERG polymorphisms are associated with cLQTS, tumors, schizophrenia, cardiovascular diseases, epilepsy, and skeletal muscle dystrophy, and studies over the past 10 years show that drugs with risk of LQTS are mainly antiarrhythmics, antihypertensives, and drugs for psychiatric disorders, epilepsy, or cancer [34]. In-depth study of the relationships between HERG potassium channel polymorphisms and diseases or adverse drug reactions can provide new insights into disease prevention and treatment. It can also hopefully present new targets for drug development and offer a new platform for re-evaluating drugs that have been withdrawn from the market due to severe adverse effects. This is important as such drugs might still have a lot to offer us and be valuable in a wide range of clinical applications [34].

33.6.3.1 HERG in the Clinical Application of New Drugs

Studies have shown that HERG blocking may be related to the treatment effect of the drugs. For example, Parkman et al. studied 48 patients with gastroparesis and found that the *KCNH2* polymorphism (rs3815459), an allele, correlated with the therapeutic effect of domperidone in treating this condition [34]. This suggests that proper HERG blocking may be another important approach for gastroparesis treatment.

Pharmacogenomic studies on HERG provide warnings during early-stage drug development, helping to avoid the withdrawal of useful compounds only because of their adverse effects on patients with genetic defects. To some extent, such studies may play an important role in providing more information for drug structural analysis. For example, erythromycin shows a blocking effect on HERG with an IC_{50} of approximately 59.3 uM; the Tyr652Ala mutation does not affect HERG sensitivity to erythromycin, whereas the Phe656Ala substitution does so slightly, suggesting that the blocking of HERG by erythromycin is due to its size [34].

New opportunities for the clinical application of drugs can be pieced together by pharmacogenomic studies, offering a basis for safety evaluation and individualized use [34]. Studies have shown that functional deficits can be repaired by HERG channel blockers, and that structurally different HERG blockers can selectively repair functional deficits caused by mutations at different sites (e.g., fexofenadine) [34]. However, the mechanism of these repair functions remains unclear. It is possible that HERG blockers function by repairing the unstable conformations caused by mutations. This suggests that HERG channel blockers can be used to treat cLQTS. However, such treatment should be based on an individualized diagnosis and related pharmacogenomic assays to ensure that it is safe and effective. Investigations into the relationship between HERG channel polymorphisms and drug sensitivity have all been performed *in vitro* [34].

Many studies have explored the effect of Tyr652Ala on drug sensitivity and have found that this amino acid mutation significantly decreases the blocking effects of most drugs, and it may reduce the risks of adverse effects for drugs that can induce QT interval abnormality. Drugs with LQTS risks may be safe and effective in populations that carry this site mutation. The Tyr652Ala polymorphism shows clear differences in racial and gender distribution. Overall, its clinical significance certainly merits further investigation [34].

The relationships between CYP2B6 and *KCNH2* polymorphisms and the effects of methadone are yet to be investigated in detail, but will be of important clinical value [34]. Eap et al. [34,95] studied 179 subjects and found that, in CYP2B6*6 carriers vs noncarriers, the risk of methadone-induced QT interval prolongation was increased 4.5-fold. *In vitro* studies have found that the blocking effect of S-methadone on HERG is 3.5-fold that of R-methadone, but there has been no research into the relationship between *KCNH2* polymorphisms and the HERG effects of these two methadone isomers. Some *in vitro* studies have suggested that *KCNH2* polymorphisms affect channel drug sensitivity and enhance the risk of QT abnormality syndromes; however, this needs to be verified with large clinical trials.

In addition, amino acid mutations at position 652 of the HERG potassium channel may not always affect HERG blocker sensitivity [34]. For example, mutations at this position have been found to increase the blocking effect of capsaicin on HERG I_{Kr} current, decreasing the IC_{50} to about a quarter of the value for the wild- type channel. Contrary to this, amino acid mutations at position 656 showed no difference compared with wild type [34]. The physiological functions of HERG potassium channels, their role in regulating disease occurrence and development, and their relationship with adverse as well as treatment effects of drugs merits further studies.

Pharmacogenomic studies on HERG potassium channels can provide a new basis for reevaluating drugs that are of value in a wide range of clinical applications but have been withdrawn merely because of their rarely occurring adverse effects [34]. Such studies can also offer new insight into the development of individualized clinical treatment services. Although at the present time the results of such studies have not been clinically applied, they are of great importance in individualized applications of drugs carrying QT abnormality risks, in evaluating the safety of withdrawn drugs, and those during early-stage development [34].

33.6.4 Rare Mutations

With the identification of rare mutations in ion channel genes underlying cLQTS and with the investigation of large pedigrees, the phenomenon of incomplete penetrance has become apparent, as previously discussed [6,23]. In a given pedigree, people with a LQTS mutation but normal QT intervals at baseline can be described as "subclinical" and may experience TdP after exposure to a QT-prolonging drug. When cLQTS disease genes are screened for, mutations are identified in ~10% of diLQTS subjects [73].

33.6.5 Common Variants and Common Polymorphisms in Ion Channel Genes

More common polymorphisms in ion channel and other genes have been described that have minor functional effects at baseline yet, when compounded with a QT-prolonging drug, may increase TdP risk. In one study of patients treated with dofetilide, the common *SCN5A* polymorphism H558R was over-represented in those with TdP (37.5%) compared with those without (14.4%) [6,96]. Other common ion channel polymorphisms implicated in small diLQTS studies include *KCNH2* R1047L and *KCNE2* T8A and Q9E [97]. A survey of a large collection of diLTQS patients compared with drug-exposed patients with minimal QT prolongation and with population controls revealed that the polymorphism D85N in *KCNE1* (an important subunit for I_{Ks} function) was a risk factor [98]. The allele frequency of D85N was ~8% among subjects with diTdP and 1–2% in controls [5].

Most variants studied to date have been "nonsynonymous" and data comparing the function of wild-type channels to that of variant channels can often be generated *in vitro* [5]. One variant that is relatively common in African subjects, with a minor allele frequency of ~13%, results in S1103Y (substitution of tyrosine for serine at position 1103 in the cardiac sodium channel). However, this variant is unusual in subjects of other ethnicities [99].

33.6.6 Genome-Wide Association Studies

Genome-wide association studies (GWAS) have been used to study variability in the baseline QT interval [100–102]. Such studies encompassed tens of thousands of individuals, identified multiple loci at which common variation influences variability in the QT interval [5]. Some of these are in ion channel genes, suggesting that rare variants in these genes can cause diseases like cLQTS, whereas common variants modulate the baseline QT interval. The QTc interval is known to be influenced by genetic factors and heritability estimates are between 25% and 52% [103–105]. In the Twins UK study [106], a UK-based sample of mostly female twins of European ancestry, the proportions of additive genetic influences have been estimated as 55% for resting heart rate, 60% for uncorrected QT interval and 50% for QTc.

As Jamshidi et al. [105] describe, until recently, research into genetic factors influencing QT interval was limited to candidate genes known to have a role in arrhythmogenesis. This was on the basis of their involvement in the congenital monogenic diseases LQTS and short QT syndrome [105,107,108]. However, advances in biotechnology have now made genome-wide association studies (GWAS) possible. GWAS have the potential to identify loci that had not previously been targeted as having a role in the trait or disease, thereby highlighting potentially novel biological pathways [109]. This is in contrast to candidate gene studies in which genes are selected on the basis of known or suspected disease mechanisms. An early GWAS for the QTc interval [100] based on selection of individuals from the extreme tails of the population-based QTc interval distribution identified a common variant in the nitric oxide synthase 1 adaptor protein (*NOS1AP*) gene region, and this was consistently confirmed in later studies [101,102,110–113]. The *NOS1AP* variant has been estimated to explain up to only 1.5% of QTc variance [100,105] suggesting the need for more studies with the potential to detect additional common genetic variants, which are likely to be of modest effect size.

Recent advances include meta-analyses of GWAS investigating QT interval duration in population-based cohorts by a number of consortia [101,102]. The QTGEN and QTSCD [101,102] consortia found that common variants in a number of genes previously known to cause congenital LQTS (*KCNQ1*, *KCNH2*, *KCNE1*, and *KCNJ2*) and short QT syndrome (*SCN5A*) were among the most strongly associated with the QT interval in these population-based cohorts. These studies have contributed many newly associated loci to this complex trait, and have suggested a cumulative effect of individual variants on QT interval [105]. Two of the novel loci contained genes with established electrophysiological function (*ATP1B1* and *PLN*) and a third locus on 16q21 was near the genes *GINS3* and *NDRG4*. These genes have been associated with myocardial repolarization in zebrafish experiments [105].

Indeed, one of the strongest signals to emerge from GWAS analysis of the baseline QT interval is variation at the nitric oxide synthase 1 adaptor protein (*NOS1AP*) locus on chromosome 1, which encodes a protein for neuronal nitric oxide synthase. *In vitro* and *in vivo* studies suggest that it modulates both potassium and calcium current function [5,114]. Following from this is the investigation of the role of QT-modulating genes in diTdP and other traits. There is some evidence that *NOS1AP* variants can act as modulators of arrhythmia risk in cLQTS [115,116] and are implicated in sudden death in the general population [117]. Further evidence for the role of the *NOS1AP* locus in diTdP was provided by a study that found a 5-fold increase in diTdP risk in patients in whom amiodarone was the culprit drug [5,118].

33.7 FUTURE DIRECTIONS

33.7.1 Screening, Stratification, Personalization

Despite substantial advances in the field since potentially proarrhythmic drugs were first described, it has become clear that there are several factors which contribute to an individual's personal risk of diLQTS and TdP. The known clinical risk factors for acquired and drug-induced TdP can serve to separate patients into high- or low-risk subgroups for development of this life-threatening arrhythmia; however, they cannot identify risk on an individual level.

For more accurate risk stratification, the genomic contribution to drug-induced TdP risk must be further evaluated and refined, as it appears that here lies more of the unpredictability of this condition. Indeed, the fact that small studies have uncovered first-degree relatives of patients with diTdP who themselves have a prolonged QTc interval on exposure to a QT-prolonging drug [119] supports the role of a genetic component in risk.

In the future, additional coding and noncoding modifier variants are likely to be discovered as novel technologies, such as exome and whole genome sequencing, are used to genotype LQTS patients. The incorporation of these variants into risk stratification algorithms is likely to have a significant impact on LQTS management [120]. Other future directions for LQTS studies include the functional effects of mutations and their correlation with risk—for example, using induced pluripotent stem cells (iPSC) derived directly from patient samples.

The potential role of rare variants in diTdP can be further examined by some of the newer approaches to sequencing [5]. In comparison with background rates ascertained by the 1000 Genomes and the Exome Sequencing projects a recent study [72] sequenced 79 candidate arrhythmia genes in a small number of subjects with diLQTS and identified a high proportion of individuals with rare variants. Ascertaining the role of genetic variation in diTdP will require gathering together a large number of well-phenotyped cases across different ethnic groups and will cover several different drugs and drug classes.

It is a realistic prospect that, in the future, we will include clinically relevant information on pharmacogenetic variation in patient records. With the decreasing costs of genotyping and sequencing, future limiting factors, other than cost, are more likely to be data interpretation and wider ethical issues resulting from the identification of variants of uncertain clinical significance. The turnaround time for results will also have to be rapid and reliable, but with advice limited in some cases to enhanced risk than a definitive "yes" or "no." Thus, some clinicians may be less inclined to support adoption of pharmacogenomic testing for diTdP on a wide scale.

Pharmacogenetic data may influence the clinical management of patients in the future—for example, by focusing on those who should be monitored more closely on treatment, those in whom drug dosage should be altered, or those for whom a particular therapy should be avoided altogether. However, before such genotype-based healthcare can be implemented, several significant issues must be addressed such as: the strength of the evidence associating genetic variation with variable drug outcomes; interaction with nongenetic risk; variable effects of genetic variants across ethnicities; reliability of genotyping assays; the development of electronic medical-records systems; and the attitudes of both patients and their clinicians [5].

33.7.2 Proposed Methods of Risk Stratification

Short-term QT interval variability (as measured from 30 consecutive beats) is increased in patients with a history of drug-induced long-QT syndrome, suggesting that this variability could prove to be a useful noninvasive, easily obtainable parameter for aiding the identification of patient at risk for potentially life-threatening arrhythmia in the context of drugs with QT-prolonging potential [121].

The development of a torsadogenic score to assess proarrhythmic drug risk would be most beneficial and this is the ultimate goal of the ARITMO (Arrhythmogenic Potential of Drugs) project [122]. In this study, the FDA Adverse Event Reporting System (FAERS) is used to detect signals of increased torsadogenic risk. The population is large (including all US reports and serious/unexpected adverse drug reactions from other countries) and this feature makes the FAERS an incredibly useful resource for exploring more rare adverse drug reactions (ADRs) such as TdP [122]. FAERS's initial findings will need to be further validated by integrating data from other sources of clinical data. However, it is important to bear in mind the limitations of current methods and to remind ourselves that, for more recently marketed drugs, careful monitoring should be maintained. It appears that *in vitro/in vivo* assays and thorough QT studies strengthen the role of postmarketing surveillance in early and timely detection of safety signals [122].

33.7.3 Unanswered Questions and Priority Areas for Research

Understanding the nature and extent of variability among human genomes will allow prediction of individual susceptibility to disease, adverse drug reactions, and even beneficial drug effects. Achieving this vision of personalized medicine requires in-depth understanding of the number of variants in specific candidate genes and across the genome, the way in which these variants affect individual gene function, and the ways in which they may be analyzable. There are enormous challenges in executing this vision. These include the requirement for new technologies for rapid genotyping; the development of approaches to organization and analysis of very large data sets; associated ethical issues; and cost and commercialization [4].

These challenges are balanced by a set of opportunities including the true possibility of delivering improved care and developing new therapies through a better understanding biology. The key first step in this research is clear definition and identification of target phenotypes, so it is essential that clinicians be educated to become leaders in this new field.

In considering priority areas for future research, we must consider not only what would happen in an ideal world but also what might realistically be readily absorbed into the general clinical workflow and be ethically acceptable. As we look to the future, we must address significant issues before genotype-based healthcare can be implemented. Most important is how to analyze the huge amounts of data generated and how to deal with that information. Ethical considerations, which even now are the subject of heated worldwide debate, must not be underestimated.

With advances in personal genomics, high-accuracy risk prediction for adverse drug effects is a real prospect. Understanding the mechanisms underlying drug actions, especially toxicities such as diLQTS, remains the best way forward in choosing the most appropriate therapies for individual patients, and in developing new, safe, and effective treatments.

GLOSSARY

Andersen-Tawil syndrome	a condition characterized by episodic paralysis, ventricular arrhythmias, and prolonged QT interval
Cascade screening	method of ascertaining gene carriers by testing the patient's close family members
Copy number variant	form of DNA variation in which a certain sequence is present in differing numbers of copies, and is different across individuals. Many CNVs are nonpathogenic common variants
Genome-wide association study (GWAS)	a study in which SNPs are tested for association with a disease in a case-control design
Monogenic	pertaining to a disorder occurring as a direct consequence of single gene defects
Penetrance	probability of a characteristic being present given a certain genotype
Phenotype	measurable characteristics of an organism. These may derive from the genotype, the environment, or their combination
Single nucleotide polymorphism (SNP)	any polymorphic variation at a single nucleotide
Syncope	transient loss of consciousness

Timothy syndrome rare childhood condition that leads to cardiac arrhythmia, a spectrum of neurological and developmental abnormalities, and syndactyly

Torsades de Pointes (TdP) distinctive form of polymorphic ventricular tachycardia characterized by twisting of the QRS complexes around the isoelectric line. It is associated with QT prolongation

Ventricular fibrillation (VF) rapid, chaotic, nonrepetitive waveform in which there is uncoordinated contraction of the cardiac muscle of the ventricles. VF is clinically associated with loss of effective blood circulation and, if not immediately treated, death

Ventricular tachycardia (VT) sequence of three or more beats in a row, with wide QRS complex (QRS >120 ms) at a ventricular rate exceeding 100 beats/min

ACRONYMS AND ABBREVIATIONS

GWAS	genome-wide association study
NME	new molecular entity
SNP	Single-nucleotide polymorphism
ECG	electrocardiogram
APD	action potential duration
EAD	early after depolarization
I_{Kr}	rapid delayed rectifier potassium current
I_{Ks}	slow delayed rectifier potassium current
KCNQ1	potassium voltage-gated channel, KQT-like subfamily, member 1
RWS	Romano-Ward syndrome
JLNS	Jervell and Lange-Nielsen syndrome
KCNH2	potassium voltage-gated channel, subfamily H, member 2
SCN5A	sodium voltage-gated channel, type V, α-subunit
ANK2	ankyrin 2
KCNE1	potassium voltage-gated channel, I_{SK}-related subfamily, member1
KCNE2	potassium voltage-gated channel, I_{SK}-related subfamily, member 2
KCNJ2	potassium channel, inwardly rectifying, subfamily J, member 2
CACNA1C	calcium voltage-dependent channel, L-type, α-1C subunit
CAV3	caveolin 3
SCN4B	sodium voltage-gated channel, type IV, β subunit
AKAP9	A-kinase anchor protein 9
SNTA1	syntrophin α1
KCNJ5	potassium channel, inwardly rectifying, subfamily J, member 5

REFERENCES

[1] Hampton JR. The ECG made easy. 4th ed. Churchill Livingstone; 2003.

[2] Postema PG, De Jong JS, Van der Bilt IA, Wilde AA. Accurate electrocardiographic assessment of the QT interval: teach the tangent. Heart Rhythm 2008;5(7):1015–8.

[3] Ponte ML, Keller GA, Di GG. Mechanisms of drug induced QT interval prolongation. Curr Drug Saf 2010;5(1):44–53.

[4] Roden DM. Long QT syndrome: reduced repolarization reserve and the genetic link. J Intern Med 2006;259(1):59–69.

[5] Behr ER, Roden D. Drug-induced arrhythmia: pharmacogenomic prescribing? Eur Heart J 2013;34(2):89–95.

[6] Kannankeril P, Roden DM, Darbar D. Drug-induced long QT syndrome. Pharmacol Rev 2010;62(4):760–81.

[7] Camm AJ, Yap YG, Malik M. Acquired long QT syndrome. Blackwell Futura; 2004.

[8] Haverkamp W, Breithardt G, Camm AJ, Janse MJ, Rosen MR, Antzelevitch C, et al. The potential for QT prolongation and proarrhythmia by non-anti-arrhythmic drugs: clinical and regulatory implications. Report on a policy conference of the European society of cardiology. Cardiovasc Res 2000;47(2):219–33.

[9] Chiladakis J, Kalogeropoulos A, Arvanitis P, Koutsogiannis N, Zagli F, Alexopoulos D. Preferred QT correction formula for the assessment of drug-induced QT interval prolongation. J Cardiovasc Electrophysiol 2010;21(8):905–13.

[10] Goldenberg I, Moss AJ. Long QT syndrome. J Am Coll Cardiol 2008;51(24):2291–300.

[11] Sagie A, Larson MG, Goldberg RJ, Bengtson JR, Levy D. An improved method for adjusting the QT interval for heart rate (the Framingham Heart Study). Am J Cardiol 1992;70(7):797–801.

[12] Luo S, Michler K, Johnston P, Macfarlane PW. A comparison of commonly used QT correction formulae: the effect of heart rate on the QTc of normal ECGs. J Electrocardiol 2004;37(Suppl.):81–90.

[13] Grant AO. Cardiac ion channels. Circ Arrhythm Electrophysiol 2009;2(2):185–94.

[14] Roden DM, Viswanathan PC. Genetics of acquired long QT syndrome. J Clin Invest 2005;115(8):2025–32.

[15] Abrams DJ, Perkin MA, Skinner JR. Long QT syndrome. BMJ 2010;340:b4815.

[16] Crotti L, Celano G, Dagradi F, Schwartz PJ. Congenital long QT syndrome. Orphanet J Rare Dis 2008;3:18.

[17] Splawski I, Shen J, Timothy KW, Lehmann MH, Priori S, Robinson JL, et al. Spectrum of mutations in long-QT syndrome genes. KVLQT1, HERG, SCN5A, KCNE1, and KCNE2. Circulation 2000;102(10):1178–85.

[18] Schwartz PJ, Priori SG, Spazzolini C, Moss AJ, Vincent GM, Napolitano C, et al. Genotype-phenotype correlation in the long-QT syndrome: gene-specific triggers for life-threatening arrhythmias. Circulation 2001;103(1):89–95.

[19] Moss AJ, Goldenberg I. Importance of knowing the genotype and the specific mutation when managing patients with long-QT syndrome. Circ Arrhythm Electrophysiol 2008;1:219–26.

[20] Schwartz PJ. Idiopathic long QT syndrome: progress and questions. Am Heart J 1985;109(2):399–411.

[21] Schwartz PJ, Moss AJ, Vincent GM, Crampton RS. Diagnostic criteria for the long QT syndrome. An update. Circulation 1993;88(2):782–4.

[22] Hofman N, Wilde AAM, Kaab S, van Langen IM, Tanck MWT, Mannens MMAM, et al. Diagnostic criteria for congenital long QT syndrome in the era of molecular genetics: do we need a scoring system? Eur Heart J 2007;28:575–80.

[23] Priori SG, Napolitano C, Schwartz PJ. Low penetrance in the long-QT syndrome: clinical impact. Circulation 1999;99(4):529–33.

[24] Mullally J, Goldenberg I, Moss AJ, Lopes CM, Ackerman MJ, Zareba W, et al. Risk of life-threatening cardiac events among patients with long QT syndrome and multiple mutations. Heart Rhythm 2013;10(3):378–82.

[25] Moss AJ, Schwartz PJ, Crampton RS, Tzivoni D, Locati EH, MacCluer J, et al. The long QT syndrome. Prospective longitudinal study of 328 families. Circulation 1991;84(3):1136–44.

[26] Zareba W, Moss AJ, Schwartz PJ, Vincent GM, Robinson JL, Priori SG, et al. Influence of genotype on the clinical course of the long-QT syndrome. International long-QT syndrome registry research group. N Engl J Med 1998;339(14):960–5.

[27] Priori SG, Schwartz PJ, Napolitano C, Bloise R, Ronchetti E, Grillo M, et al. Risk stratification in the long-QT syndrome. N Engl J Med 2003;348(19):1866–74.

[28] Tester DJ, Ackerman MJ. Postmortem long QT syndrome genetic testing for sudden unexplained death in the young. J Am Coll Cardiol 2007;1(49):240–6.

[29] Moss AJ, Schwartz PJ, Crampton RS, Locati E, Carleen E. The long QT syndrome: a prospective international study. Circulation 1985;71(1):17–21.

[30] Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med 2004;1(350):1013–22.

[31] Selzer A, Wray HW. Quinidine syncope. Paroxysmal ventricular fibrillation occurring during treatment of chronic atrial arrhythmias. Circulation 1964;30:17–26.

[32] Morissette P, Hreiche R, Turgeon J. Drug-induced long QT syndrome and torsade de pointes. Can J Cardiol 2005;21(10):857–64.

[33] Behr ER, January C, Schulze-Bahr E, Grace AA, Kaab S, Fiszman M, et al. The International Serious Adverse Events Consortium (iSAEC) phenotype standardization project for drug-induced torsades de pointes. Eur Heart J 2013;34(26):1958–63.

[34] He FZ, McLeod HL, Zhang W. Current pharmacogenomic studies on hERG potassium channels. Trends Mol Med 2013;19(4):227–38.

[35] Lazzara R. Amiodarone and torsade de pointes. Ann Intern Med 1989;111(7):549–51.

[36] Kannankeril PJ, Roden DM. Drug-induced long QT and torsade de pointes: recent advances. Curr Opin Cardiol 2007;22(1):39–43.

[37] Shimizu W, Antzelevitch C. Cellular and ionic basis for T-wave alternans under long-QT conditions. Circulation 1999;99(11):1499–507.

[38] Passman R, Kadish A. Polymorphic ventricular tachycardia, long Q-T syndrome, and torsades de pointes. Med Clin North Am 2001;85(2):321–41.

[39] Shimizu W, Ohe T, Kurita T, Kawade M, Arakaki Y, Aihara N, et al. Effects of verapamil and propranolol on early afterdepolarizations and ventricular arrhythmias induced by epinephrine in congenital long QT syndrome. J Am Coll Cardiol 1995;26(5):1299–309.

[40] Jenzer HR, Hagemeijer F. Quinidine syncope: torsade de pointes with low quinidine plasma concentrations. Eur J Cardiol 1976;4(4):447–51.

[41] Snyders J, Knoth KM, Roberds SL, Tamkun MM. Time-, voltage-, and state-dependent block by quinidine of a cloned human cardiac potassium channel. Mol Pharmacol 1992;41(2):322–30.

[42] Roden DM, Hoffman BF. Action potential prolongation and induction of abnormal automaticity by low quinidine concentrations in canine Purkinje fibers. Relationship to potassium and cycle length. Circ Res 1985;56(6):857–67.

[43] Balser JR, Roden DM, Bennett PB. Single inward rectifier potassium channels in guinea pig ventricular myocytes. Effects of quinidine. Biophys J 1991;59(1):150–61.

[44] Antzelevitch C, Shimizu W, Yan GX, Sicouri S, Weissenburger J, Nesterenko VV, et al. The M cell: its contribution to the ECG and to normal and abnormal electrical function of the heart. J Cardiovasc Electrophysiol 1999;10(8):1124–52.

[45] Kiehn J, Thomas D, Karle CA, Schols W, Kubler W. Inhibitory effects of the class III antiarrhythmic drug amiodarone on cloned HERG potassium channels. Naunyn Schmiedebergs Arch Pharmacol 1999;359(3):212–9.

[46] Kamiya K, Nishiyama A, Yasui K, Hojo M, Sanguinetti MC, Kodama I. Short- and long-term effects of amiodarone on the two components of cardiac delayed rectifier K(+) current. Circulation 2001;103(9):1317–24.

[47] Roden DM, Bennett PB, Snyders DJ, Balser JR, Hondeghem LM. Quinidine delays IK activation in guinea pig ventricular myocytes. Circ Res 1988;62(5):1055–8.

[48] Imaizumi Y, Giles WR. Quinidine-induced inhibition of transient outward current in cardiac muscle. Am J Physiol 1987;253(3 Pt 2):H704–8.

[49] Sanguinetti MC, Jurkiewicz NK. Two components of cardiac delayed rectifier K+ current. Differential sensitivity to block by class III antiarrhythmic agents. J Gen Physiol 1990;96(1):195–215.

[50] Sanguinetti MC, Jurkiewicz NK. Delayed rectifier outward K+ current is composed of two currents in guinea pig atrial cells. Am J Physiol 1991;260(2 Pt 2):H393–9.

[51] Makkar RR, Fromm BS, Steinman RT, Meissner MD, Lehmann MH. Female gender as a risk factor for torsades de pointes associated with cardiovascular drugs. JAMA 1993;270(21):2590–7.

[52] Drolet B, Vincent F, Rail J, Chahine M, Deschenes D, Nadeau S, et al. Thioridazine lengthens repolarization of cardiac ventricular myocytes by blocking the delayed rectifier potassium current. J Pharmacol Exp Ther 1999;288(3):1261–8.

[53] Nachimuthu S, Assar MD, Schussler JM. Drug-induced QT interval prolongation: mechanisms and clinical management. Ther Adv Drug Safety 2012;3(5):241–53.

[54] Roden DM, Woosley RL, Primm RK. Incidence and clinical features of the quinidine-associated long QT syndrome: implications for patient care. Am Heart J 1986;111(6):1088–93.

[55] Tamargo J. Drug-induced torsade de pointes: from molecular biology to bedside. Jpn J Pharmacol 2000;83(1):1–19.

[56] Olshansky B, Martins J, Hunt S. N-acetyl procainamide causing torsades de pointes. Am J Cardiol 1982;50(6):1439–41.

[57] Lehmann MH, Hardy S, Archibald D, Quart B, MacNeil DJ. Sex difference in risk of torsade de pointes with d, l-sotalol. Circulation 1996;94(10):2535–41.

[58] Stambler BS, Wood MA, Ellenbogen KA, Perry KT, Wakefield LK, VanderLugt JT. Efficacy and safety of repeated intravenous doses of ibutilide for rapid conversion of atrial flutter or fibrillation. Ibutilide repeat dose study investigators. Circulation 1996;94(7):1613–21.

[59] Davila I, Sastre J, Bartra J, del CA, Jauregui I, Montoro J, et al. Effect of H1 antihistamines upon the cardiovascular system. J Investig Allergol Clin Immunol 2006;16(Suppl 1):13–23.

[60] Vieweg WV. New generation antipsychotic drugs and QTc interval prolongation. Prim Care Companion J Clin Psychiatry 2003;5(5):205–15.

[61] Ray WA, Meredith S, Thapa PB, Hall K, Murray KT. Cyclic antidepressants and the risk of sudden cardiac death. Clin Pharmacol Ther 2004;75(3):234–41.

[62] Anderson ME, Mazur A, Yang T, Roden DM. Potassium current antagonist properties and proarrhythmic consequences of quinolone antibiotics. J Pharmacol Exp Ther 2001;296(3):806–10.

[63] Huang BH, Wu CH, Hsia CP, Yin CC. Azithromycin-induced torsade de pointes. Pacing Clin Electrophysiol 2007;30(12):1579–82.

[64] Wesche DL, Schuster BG, Wang WX, Woosley RL. Mechanism of cardiotoxicity of halofantrine. Clin Pharmacol Ther 2000;67(5):521–9.

[65] Eisenhauer MD, Eliasson AH, Taylor AJ, Coyne Jr PE, Wortham DC. Incidence of cardiac arrhythmias during intravenous pentamidine therapy in HIV-infected patients. Chest 1994;105(2):389–95.

[66] Wysowski DK, Corken A, Gallo-Torres H, Talarico L, Rodriguez EM. Postmarketing reports of QT prolongation and ventricular arrhythmia in association with cisapride and Food and Drug Administration regulatory actions. Am J Gastroenterol 2001;96(6):1698–703.

[67] Aerssens J, Paulussen AD. Pharmacogenomics and acquired long QT syndrome. Pharmacogenomics 2005;6(3):259–70.

[68] Drew BJ, Ackerman MJ, Funk M, Gibler WB, Kligfield P, Menon V, et al. Prevention of torsade de pointes in hospital settings: a scientific statement from the American Heart Association and the American College of Cardiology Foundation. J Am Coll Cardiol 2010;55(9):934–47.

[69] Stringer J, Welsh C, Tommasello A. Methadone-associated Q-T interval prolongation and torsades de pointes. Am J Health Syst Pharm 2009;66(9):825–33.

[70] Krantz MJ, Martin J, Stimmel B, Mehta D, Haigney MC. QTc interval screening in methadone treatment. Ann Intern Med 2009;150(6):387–95.

[71] Napolitano C, Schwartz PJ, Brown AM, Ronchetti E, Bianchi L, Pinnavaia A, et al. Evidence for a cardiac ion channel mutation underlying drug-induced QT prolongation and life-threatening arrhythmias. J Cardiovasc Electrophysiol 2000;11(6):691–6.

[72] Ramirez AH, Shaffer CM, Delaney JT, Sexton DP, Levy SE, Rieder MJ, et al. Novel rare variants in congenital cardiac arrhythmia genes are frequent in drug-induced torsades de pointes. Pharmacogenomics J 2013;13(4):325–9.

[73] Paulussen AD, Gilissen RA, Armstrong M, Doevendans PA, Verhasselt P, Smeets HJ, et al. Genetic variations of KCNQ1, KCNH2, SCN5A, KCNE1, and KCNE2 in drug-induced long QT syndrome patients. J Mol Med (Berl) 2004;82(3):182–8.

[74] Chiang C-E. Drug-induced long QT syndrome. J Med Biol Eng 2006;26(3):107–13.

[75] Mihic A, Chauhan VS, Gao X, Oudit GY, Tsushima RG. Trafficking defect and proteasomal degradation contribute to the phenotype of a novel KCNH2 long QT syndrome mutation. PLoS One 2011;6(3):e18273.

[76] Aidery P, Kisselbach J, Gaspar H, Baldea I, Schweizer PA, Becker R, et al. Identification and functional characterization of the novel human ether-a-go-go-related gene (hERG) R744P mutant associated with hereditary long QT syndrome 2. Biochem Biophys Res Commun 2012;418(4):830–5.

[77] Krishnan Y, Zheng R, Walsh C, Tang Y, McDonald TV. Partially dominant mutant channel defect corresponding with intermediate LQT2 phenotype. Pacing Clin Electrophysiol 2012;35(1):3–16.

[78] Barc J, Briec F, Schmitt S, Kyndt F, Le CM, Baron E, et al. Screening for copy number variation in genes associated with the long QT syndrome: clinical relevance. J Am Coll Cardiol 2011;57(1):40–7.

[79] Levine TB, Bernink PJ, Caspi A, Elkayam U, Geltman EM, Greenberg B, et al. Effect of mibefradil, a T-type calcium channel blocker, on morbidity and mortality in moderate to severe congestive heart failure: the MACH-1 study. Mortality assessment in congestive heart failure trial. Circulation 2000;101(7):758–64.

[80] Richichi I, Valsecchi O, Falcone C, Sofi A, Zanini L. A case of "torsade de pointe" during vincamine therapy. Minerva Cardioangiol 1978;26(3):197–200.

[81] Monahan BP, Ferguson CL, Killeavy ES, Lloyd BK, Troy J, Cantilena Jr LR. Torsades de pointes occurring in association with terfenadine use. JAMA 1990;264(21):2788–90.

[82] Woosley RL, Chen Y, Freiman JP, Gillis RA. Mechanism of the cardiotoxic actions of terfenadine. JAMA 1993;269(12):1532–6.

[83] De PF, Poluzzi E, Montanaro N. QT-interval prolongation by non-cardiac drugs: lessons to be learned from recent experience. Eur J Clin Pharmacol 2000;56(1):1–18.

[84] Yap YG, Camm J. Risk of torsades de pointes with non-cardiac drugs. Doctors need to be aware that many drugs can cause qt prolongation. BMJ 2000;320(7243):1158–9.

[85] Temple RJ, Himmel MH. Safety of newly approved drugs: implications for prescribing. JAMA 2002;287(17):2273–5.

[86] Bloomfield DM, Kost JT, Ghosh K, Hreniuk D, Hickey LA, Guitierrez MJ, et al. The effect of moxifloxacin on QTc and implications for the design of thorough QT studies. Clin Pharmacol Ther 2008;84(4):475–80.

[87] Fosser C, Duczynski G, Agin M, Wicker P, Darpo B. Comparison of manual and automated measurements of the QT interval in healthy volunteers: an analysis of five thorough QT studies. Clin Pharmacol Ther 2009;86(5):503–6.

[88] Darpo B. The thorough QT/QTc study 4 years after the implementation of the ICH E14 guidance. Br J Pharmacol 2010;159(1):49–57.

[89] Lahti AL, Kujala VJ, Chapman H, Koivisto AP, Pekkanen-Mattila M, Kerkela E, et al. Model for long QT syndrome type 2 using human iPS cells demonstrates arrhythmogenic characteristics in cell culture. Dis Model Mech 2012;5(2):220–30.

[90] Harley CA, Jesus CS, Carvalho R, Brito RM, Morais-Cabral JH. Changes in channel trafficking and protein stability caused by LQT2 mutations in the PAS domain of the HERG channel. PLoS One 2012;7(3):e32654.

[91] Delisle BP, Anderson CL, Balijepalli RC, Anson BD, Kamp TJ, January CT. Thapsigargin selectively rescues the trafficking defective LQT2 channels G601S and F805C. J Biol Chem 2003;278(37):35749–54.

[92] Yao Y, Teng S, Li N, Zhang Y, Boyden PA, Pu J. Aminoglycoside antibiotics restore functional expression of truncated HERG channels produced by nonsense mutations. Heart Rhythm 2009;6(4):553–60.

[93] Rajamani S, Eckhardt LL, Valdivia CR, Klemens CA, Gillman BM, Anderson CL, et al. Drug-induced long QT syndrome: hERG K+ channel block and disruption of protein trafficking by fluoxetine and norfluoxetine. Br J Pharmacol 2006;149(5):481–9.

[94] Gong Q, Stump MR, Zhou Z. Inhibition of nonsense-mediated mRNA decay by antisense morpholino oligonucleotides restores functional expression of hERG nonsense and frameshift mutations in long-QT syndrome. J Mol Cell Cardiol 2011;50(1):223–9.

[95] Eap CB, Crettol S, Rougier JS, Schlapfer J, Sintra GL, Deglon JJ, et al. Stereoselective block of hERG channel by (S)-methadone and QT interval prolongation in CYP2B6 slow metabolizers. Clin Pharmacol Ther 2007;81(5):719–28.

[96] Mank-Seymour AR, Richmond JL, Wood LS, Reynolds JM, Fan YT, Warnes GR, et al. Association of torsades de pointes with novel and known single nucleotide polymorphisms in long QT syndrome genes. Am Heart J 2006;152(6):1116–22.

[97] Fitzgerald PT, Ackerman MJ. Drug-induced torsades de pointes: the evolving role of pharmacogenetics. Heart Rhythm 2005;2(2 Suppl):S30–7.

[98] Kaab S, Crawford DC, Sinner MF, Behr ER, Kannankeril PJ, Wilde AA, et al. A large candidate gene survey identifies the KCNE1 D85N polymorphism as a possible modulator of drug-induced torsades de pointes. Circ Cardiovasc Genet 2012;5(1):91–9.

[99] Plant LD, Bowers PN, Liu Q, Morgan T, Zhang T, State MW, et al. A common cardiac sodium channel variant associated with sudden infant death in African Americans, SCN5A S1103Y. J Clin Invest 2006;116(2):430–5.
[100] Arking DE, Pfeufer A, Post W, Kao WH, Newton-Cheh C, Ikeda M, et al. A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. Nat Genet 2006;38(6):644–51.
[101] Newton-Cheh C, Eijgelsheim M, Rice KM, de Bakker PI, Yin X, Estrada K, et al. Common variants at ten loci influence QT interval duration in the QTGEN study. Nat Genet 2009;41(4):399–406.
[102] Pfeufer A, Sanna S, Arking DE, Muller M, Gateva V, Fuchsberger C, et al. Common variants at ten loci modulate the QT interval duration in the QTSCD study. Nat Genet 2009;41(4):407–14.
[103] Newton-Cheh C, Larson MG, Corey DC, Benjamin EJ, Herbert AG, Levy D, et al. QT interval is a heritable quantitative trait with evidence of linkage to chromosome 3 in a genome-wide linkage analysis: the Framingham Heart Study. Heart Rhythm 2005;2(3):277–84.
[104] Multikainen S, Ortega-Alonso A, Alen M, Kaprio J, Karjalaien J, Rantanen T, et al. Genetic influences on resting electrocradiographic variables in older women: a twin study. Ann Noninvas Electrocardiol 2009;1(14):57–64.
[105] Jamshidi Y, Nolte IM, Spector TD, Snieder H. Novel genes for QTc interval. How much heritability is explained, and how much is left to find? Genome Med 2010;2(5):35.
[106] Dalageorgou C, Ge D, Jamshidi Y, Nolte IM, Riese H, Savalieva I, et al. Heritability of QT interval: how much is explained by genes for resting heart rate? J Cardiovasc Electrophysiol 2008;19:386–91.
[107] Newton-Cheh C, Guo CY, Larson MG, Musone SL, Surti A, Camargo AL, et al. Common genetic variation in KCNH2 is associated with QT interval duration: the Framingham Heart Study. Circulation 2007;116(10):1128–36.
[108] Pfeufer A, Jalilzadeh S, Perz S, Mueller JC, Hinterseer M, Illig T, et al. Common variants in myocardial ion channel genes modify the QT interval in the general population: results from the KORA study. Circ Res 2005;96(6):693–701.
[109] Donnelly P. Progress and challenges in genome-wide association studies in humans. Nature 2008;456:728–31.
[110] Tobin MD, Kahonen M, Braund P, Nieminem T, Hajat C, Tomaszewski M, et al. Gender and effects of a common genetic variant in the NOS1 regulator NOS1AP on cardiac repolarization in 3761 individuals from two independent populations. Int J Epidemiol 2008;37:1132–41.
[111] Raitakari OT, Blom-Nyholm J, Koskinen TA, Kahonen M, Viikari JS, Lehtimaki T. Common variation in NOS1AP and KCNH2 genes and QT interval duration in young adults. The Cardiovascular Risk in Young Finns Study. Ann Med 2009;1(41):144–51.
[112] Lehtinen AB, Newton-Cheh C, Ziegler JT, Langefeld CD, Freedman BI, Daniel KR, et al. Association of NOS1AP genetic variants with QT interval duration in families from the Diabetes Heart Study. Diabetes 2008;57:1108–14.
[113] Eijgelsheim M, Aarnoudse AL, Rivandeneira F, Kors JA, Witteman JC, Hofman A, et al. Identification of a common variant at the NOS1AP locus strongly associated to QT-interval duration. Hum Mol Genet 2009;1(18):347–57.
[114] Chang KC, Barth AS, Sasano T, Kizana E, Kashiwakura Y, Zhang Y, et al. CAPON modulates cardiac repolarization via neuronal nitric oxide synthase signaling in the heart. Proc Natl Acad Sci USA 2008;105(11):4477–82.
[115] Crotti L, Monti MC, Insolia R, Peljto A, Goosen A, Brink PA, et al. NOS1AP is a genetic modifier of the long-QT syndrome. Circulation 2009;120(17):1657–63.
[116] Tomas M, Napolitano C, De GL, Bloise R, Subirana I, Malovini A, et al. Polymorphisms in the NOS1AP gene modulate QT interval duration and risk of arrhythmias in the long QT syndrome. J Am Coll Cardiol 2010;55(24):2745–52.
[117] Kao WH, Arking DE, Post W, Rea TD, Sotoodehnia N, Prineas RJ, et al. Genetic variations in nitric oxide synthase 1 adaptor protein are associated with sudden cardiac death in US white community-based populations. Circulation 2009;119(7):940–51.
[118] Jamshidi Y, Nolte IM, Dalageorgou C, Zheng D, Johnson T, Bastiaenen R, et al. Common variation in the NOS1AP gene is associated with drug-induced QT prolongation and ventricular arrhythmia. J Am Coll Cardiol 2012;60(9):841–50.
[119] Kannankeril PJ, Roden DM, Norris KJ, Whalen SP, George Jr AL, Murray KT. Genetic susceptibility to acquired long QT syndrome: pharmacologic challenge in first-degree relatives. Heart Rhythm 2005;2(2):134–40.
[120] Mahida S, Hogarth AJ, Cowan C, Tayebjee MH, Graham LN, Pepper CB. Genetics of congenital and drug-induced long QT syndromes: current evidence and future research perspectives. J Interv Card Electrophysiol 2013;37(1):9–19.
[121] Kallergis EM, Goudis CA, Simantirakis EN, Kochiadakis GE, Vardas PE. Mechanisms, risk factors, and management of acquired long QT syndrome: a comprehensive review. Sci World J 2012;2012:212178.
[122] Poluzzi E, Raschi E, Koci A, Moretti U, Spina E, Behr ER, et al. Antipsychotics and torsadogenic risk: signals emerging from the US FDA adverse event reporting system database. Drug Saf 2013;36:467–79.

SUGGESTED WEBSITES

Genes	Genes: www.omim.org.
Torsades de Pointes	Torsades de Pointes: www.torsades.org.
QT-prolonging drugs	QT-prolonging drugs: www.qtdrugs.org.
Sudden death	Sudden death: www.sads.org.

Chapter 34

Stratified Medicine for Pancreatic Cancer

David K. Chang,[a,b,c,d,e] Nigel B. Jamieson,[a,f] Sean M. Grimmond,[a,g]
and Andrew V. Biankin[a,b,c,d,e] for the Australian Pancreatic Cancer
Genome Initiative

[a]*Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, UK,* [b]*The Kinghorn Cancer Centre, Cancer Division, Garvan Institute of Medical Research, Australia,* [c]*St Vincent's Clinical School, Faculty of Medicine, University of NSW, Australia,* [d]*Department of Surgery, Bankstown Hospital, Australia,* [e]*South Western Sydney Clinical School, Faculty of Medicine, University of NSW, Australia,* [f]*Academic Unit of Surgery, Faculty of Medicine, Glasgow Royal Infirmary, University of Glasgow, UK,* [g]*Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, University of Queensland, Australia*

Chapter Outline

34.1 Overview 807
34.2 Current Treatment Paradigm 807
34.3 Stratified Therapeutic Approach 808
  34.3.1 Metastatic Pancreatic Cancer 808
  34.3.2 Locally Advanced Pancreatic Cancer 811
  34.3.3 Resectable Pancreatic Cancer 811
34.4 Preclinical Models of Pancreatic Cancer 812
34.5 Central Knowledge Database 812
34.6 Biorepository and Privacy Issues 813
34.7 Conclusion 813
References 813

34.1 OVERVIEW

Pancreatic cancer (PC) is the fourth leading cause of cancer-related death in Western societies, including the United Kingdom, with an overall five-year survival rate of less than 5%—a figure that has not changed in 50 years [1]. Surgical resection remains the only chance of cure, with chemotherapy adding a modest benefit. With a few exceptions, most clinical trials have not shown a statistically significant or clinically meaningful survival benefit. This is perhaps not surprising, as recent genomic sequencing studies revealed that, apart from the few well-known mutations that cannot be effectively targeted, most mutations occur at a rate of less than 5%, making PC one of the most heterogeneous cancers [2,3]. This likely hampers clinical trials given that the responsive phenotype of a particular therapeutic regimen would most likely fall below the detection threshold of a conventional randomized controlled trial design.

This heterogeneity makes PC an ideal disease to explore a stratified approach through matching the right treatment to the right patient. Stratification would improve overall outcomes and quality of life by selecting the treatment most likely to be effective and minimizing unnecessary side effects from ineffective therapies. Furthermore, it would reduce the cost of ineffective drugs and their associated toxicities, as well as the cost of drug development. In 2012, three-quarters of worldwide pharmaceutical investment did not achieve the desired efficacy, resulting in high costs for healthcare providers. For example, Tarceva®, a USFDA-approved agent for PC, has a price tag of greater than US $600,000 per quality-adjusted life year.

34.2 CURRENT TREATMENT PARADIGM

Current classification and management of PC are based on morphology (Figure 34.1). After diagnosis, the patient undergoes staging investigations, followed by a stage-specific treatment pathway. Generally, patients are categorized as having resectable or unresectable disease. Patients

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00034-7
© 2014 Elsevier Inc. All rights reserved.

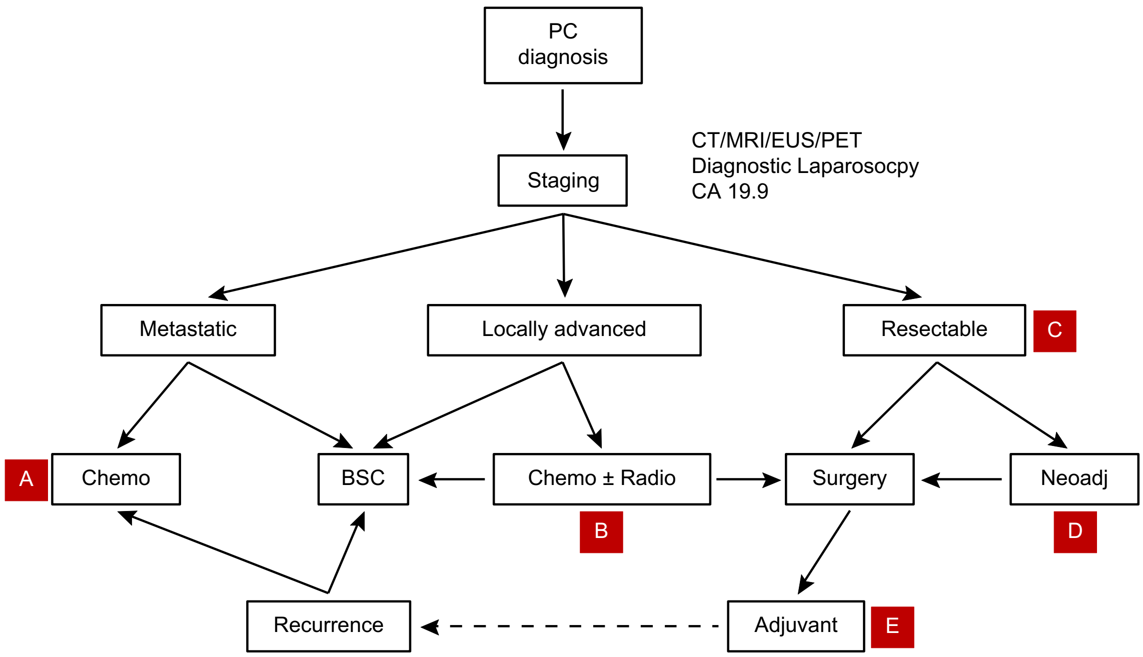


FIGURE 34.1 **Current management approach for pancreatic cancer based on morphology.** Stratification points A–E, in red, identify where biomarkers may be used to refine current management using stratification. PC—pancreatic cancer; CT—computer tomography; MRI—magnetic resonance imaging; EUS—endoscopic ultrasound; PET—positron emission tomography; BSC—best supportive care; chemo—chemotherapy; radio—radiotherapy; neoadj—neoadjuvant therapy.

with resectable PC undergo surgery or, more increasingly common, neoadjuvant chemotherapy, either as institutional policy or based on individual clinician preference. Following surgery, patients undergo chemotherapy in the adjuvant setting. For patients with unresectable PC, treatment consists of palliative chemotherapy or best supportive care depending on the individual's performance status. A wide variety of chemotherapy regimens are potentially available, including gemcitabine, 5-FU, FOLFIRINOX, and many other combinations that depend on patient fitness and clinician preferences. Radiotherapy is used sporadically in various settings.

34.3 STRATIFIED THERAPEUTIC APPROACH

34.3.1 Metastatic Pancreatic Cancer

More than 50% of patients with PC present with metastatic disease at diagnosis. The therapeutic options for these patients are either best supportive care or palliative chemotherapy. These patients are often frail as a result of cancer cachexia, and the majority are not chemotherapy candidates because of their poor performance status. Even for patients who are candidates for chemotherapy, the overall response rates have been generally disappointing, with rapidly diminishing numbers that are eligible for second- or third-line chemotherapy because of accumulating toxicity and increasing cachexia. Despite low overall response rates, palliative chemotherapy appears to be effective in an undefined subgroup of patients [4,5]. Better selection may therefore improve response rates in this subgroup and avoid ineffective chemotherapy in the remainder. This has become more relevant in recent years, with the increasing number of available chemotherapeutic options.

Better selection of patients based on predictive biomarkers of therapeutic responsiveness is the modern paradigm of clinical oncology. Through the efforts of the authors' group and others, several actionable phenotypes of therapeutic responsiveness have been proposed (Figure 34.1, stratification point A). Despite low recurrent mutation rates, it is possible to map putative actionable molecular genotypes/phenotypes in PC using a variety of tissue- and genomic-based assays. For the purpose of this review, we concentrate only on actionable molecular genotypes/phenotypes with drugs currently approved for PC or other cancer types (Table 34.1). Other emerging actionable molecular genotypes/phenotypes using repurposed or novel therapeutics are in development.

34.3.1.1 Gemcitabine

Until recently, gemcitabine was the only option for advanced PC [6]. The putative biomarkers of gemcitabine responsiveness include nucleoside transporters such as hENT1, hCNT1/3, and kinases involved in gemcitabine metabolism, such as deoxycytidine kinase (dCK) [7,8].

TABLE 34.1 PC Actionable Molecular Genotypes/Phenotypes

Actionable Genotype (Molecular Phenotype)	Therapeutic	Rationale	Molecular Characterization/ Proposed Biomarkers	Overall Prevalence
Gemcitabine responsive	Gemcitabine	Phase III clinical trial data	High *hENT1*, *hCNT1*, *hCNT3*	~14%
DDR deficient	Platinum; MMC; PARPi	Case reports; clinical trial signals FOLFIRINOX trial	*BRCA2/PALB2* mutations	8%
nab-paclitaxel responsive	*nab*-paclitaxel	Clinical trial; preclinical models	*SPARC* expression	10%
Fluoropyrimidine responsive	5-fluorouracil; capecitabine	Phase III clinical trials	Unknown	3%
Anti-EGFR responsive	Erlotinib	Phase III clinical trial data (PA3)	*KRAS* wild type	7%
Irinotecan responsive	Irinotecan	FOLFIRINOX trial	Topoisomerase 1 over expression	2%
HER2 amplified	Trastuzumab	Rescuing	*HER2* amplification	2%

Note: Data are for drugs approved for PC or other cancer type, defined using tissue- and genomic-based assays.

The most studied biomarker of therapeutic responsiveness in PC is hENT1, a membranous equilibrative nucleoside transporter encoded by the *SLC29A1* gene. There is evidence to support the role of hENT1 in gemcitabine responsiveness in PC cells both *in vitro* and *in vivo*, but its precise function as a predictive biomarker in the clinic is yet to be established. Small cohort studies and retrospective analysis of large phase III randomized control trials (RCTs), such as RTOG 9704 and ESPAC 1/3, have supported hENT1 as a predictive biomarker of adjuvant gemcitabine responsiveness [9,10]. However, a recent phase II RCT stratified by hENT1 expression, comparing gemcitabine versus CO-101 (lipophilic gemcitabine) in metastatic PC, failed to demonstrate this in metastatic disease [11]. Though the reasons for this are still unclear, the discrepancy may be due to the use of different hENT1 antibodies for immunohistochemistry; or it may be that the significance of hENT1 as a predictive biomarker is different in the metastatic as compared to the adjuvant setting; or, finally, it may be that hENT1 is not a predictive biomarker of gemcitabine responsiveness.

34.3.1.2 DNA-Damaging Agents

Perhaps the most striking cellular hallmark of cancers with mutations in the Fanconi anaemia/BRCA homologous recombination pathway is their hypersensitivity to DNA-damaging agents, such as mitomycin C and platinum, or to PARP inhibitors [12,13]. Anecdotal evidence for the potential use of DNA-damaging agents in PC that harbor these defects is mounting.

Platinum-based therapies in PC have had mixed results in clinical trials of unselected patients [14], although a recent meta-analysis of clinical trials [15] and efficacy of the FOLFIRINOX regimen [4] suggest activity in subgroups of patients. The efficacy of FOLFIRINOX in PC was demonstrated by the PRODIGE4/ACCORD 11 study [4] (overall median survival 11.1 versus 6.8 months; $p < 0.001$). FOLFIRINOX is a combination of 5-fluorouracil, irinotecan, and oxaliplatin, and can be associated with significant drug-related toxicity. Therefore, predicting responders prior to therapy could significantly improve overall outcomes. In addition, most evidence suggests that the predominant active agent in this combination is oxaliplatin, as similar response rates are seen despite dose reduction in 5-FU and irinotecan [16,17]. This implies that defining patients whose tumors harbor defects in DNA damage repair and replication (DDR) may facilitate better patient selection and remove the necessity for other drugs. While irinotecan may have synergistic activity with platinums, many consider that 5-FU adds little considering the cost of associated toxicity. Putative biomarkers of DNA-damaging agent responsiveness have not been well characterized because of complex interactions between a large number of genes involved in DNA maintenance machinery.

Platinum agents and PARP inhibitors are currently in clinical trials for the treatment of hereditary breast and ovarian cancers [13,18], recruited based on germline defects or variants in Fanconi anemia genes. However, there is increasing evidence to suggest that the responsive phenotype may be greater than just the patients harboring tumors with germline and/or somatic mutations in known DNA maintenance genes. The majority of these patients have either germline or somatic known pathogenic mutations in genes involved in DNA maintenance; however, up to 30% of them have novel mutations of unknown significance in known DNA maintenance genes or mutations in novel genes that have not been implicated in DNA maintenance.

In a collaborative effort led by the Sanger Institute to which the Australian Pancreatic Cancer Genome Initiative (APGI) contributed, an analysis of 4,942,984 mutations from 7042 cancers extracted more than 20 distinct mutational signatures [19]. The results revealed four signatures in PC that infer known biological processes: older age, *BRCA*-mediated defects in DNA damage repair, DNA mismatch repair deficiency, and a signature associated with the APOBEC family of cytidine deaminases. As a consequence, mutational signatures associated with specific defects in DNA maintenance may also be clinically useful in defining a DNA-damaging agent responsive phenotype and may have utility for patient selection in the future.

34.3.1.3 Nab-Paclitaxel

Secreted Protein Acid and Rich in Cysteine (SPARC, also known as osteonectin) regulates extracellular matrix modeling and deposition and may act as a tumor suppressor or an oncogenic driver depending on differential expression in epithelial and stromal components in different cancer types [20]. High stromal and low epithelial SPARC expression is a poor prognostic biomarker in PC [21,22]; based on its hypothesized function as an albumin "sticker," it was developed as a therapeutic target for *nab*-paclitaxel (Abraxane®) to enable "stromal depletion" and, in turn, improve drug delivery.

A phase I/II study of gemcitabine plus *nab*-paclitaxel demonstrated SPARC expression in the stroma but not in the epithelium, cosegregated with improved survival, and hence a candidate predictive biomarker for *nab*-paclitaxel responsiveness [23]. This led to the phase III MPACT (Metastatic Pancreatic Adenocarcinoma Clinical Trial) RCT, which compared gemcitabine versus gemcitabine plus Abraxane. This trial demonstrated that the addition of Abraxane conferred significant survival benefit over gemcitabine alone in patients with metastatic PC (median overall survival 8.5 versus 6.7 months; $p<0.001$) [5]; however, data concerning SPARC as a predictive biomarker are not currently available. Although the relationship between SPARC expression and *nab*-paclitaxel responsiveness is still evolving, these proof-of-concept data warrant further exploration of SPARC as a biomarker that can potentially better select patients for *nab*-paclitaxel treatment.

34.3.1.4 Erlotinib Responsiveness

In the NCIC CTG PA.3 study, the combination of erlotinib and gemcitabine demonstrated a modest but statistically significant survival advantage over gemcitabine alone in advanced PC [24]. However, in patients with Grade ≥2 skin rash, the median survival was 10.5 months—double that of those without a rash [24]. A retrospective molecular analysis of the trial did not demonstrate either *KRAS* mutation status or *EGFR* gene copy number as predictive biomarkers of erlotinib responsiveness [25]. The NCIC study was limited because tissue was available for analysis in only 32% of the patients, and even fewer patients had available *KRAS* and *EGFR* results (26% and 15%, respectively). The proportion of *KRAS* wild-type tumors (21%) was also much higher than the 7% reported in large sequencing studies using next-generation technology [2]. It remains unclear if anti-EGFR therapy can be guided by the mutation status of *KRAS*, as it is in colon cancer, where activating mutations are associated with resistance to therapy.

34.3.1.5 HER2 Amplification

HER2 amplification has been used as a predictive biomarker of trastuzumab responsiveness in breast and gastric cancer [26,27]. Preclinical PC data support the efficacy of anti-*HER2* therapy in HER2 over expressing PC [28,29]. Activation of *HER2* also transforms *HPDE* ("normal" pancreatic cell line) with activated *KRAS* and inactivated *p16/p14* and *SMAD4*, suggesting that *HER2* may be an important oncogenic driver in PC [30]. Results from trastuzumab clinical trials have been somewhat disappointing [31,32]; however, this may be a consequence of the nonstandardized assays that were used for patient selection leading to the overestimation of *HER2* amplified patients and thus the underpowering of the studies. The reported HER2 positive rates range between 0 and 80% in the literature because of different assays and cut-points. The authors' group observed only 2% *HER2* amplification (IHC 2/3+ and FISH positive) in a cohort of 469 PC patients using standardized assays performed in a national reference laboratory [33]. We also observed that *HER2*-amplified PC may be a different clinical entity that is characterized by the lack of liver metastases and the preponderance of lung and brain metastases [33].

34.3.1.6 Individualized Molecular Pancreatic Cancer Therapy Trial

In order to test the preceding actionable molecular genotypes/phenotypes, an Individualized Molecular Pancreatic Cancer Therapy (IMPaCT) trial has been established (Australian New Zealand Clinical Trial Registry ID: ACTRN12612000777897; see Figure 34.2). IMPaCT is a phase II prospective trial that randomized patients between standard therapy and stratified treatment in a first-line

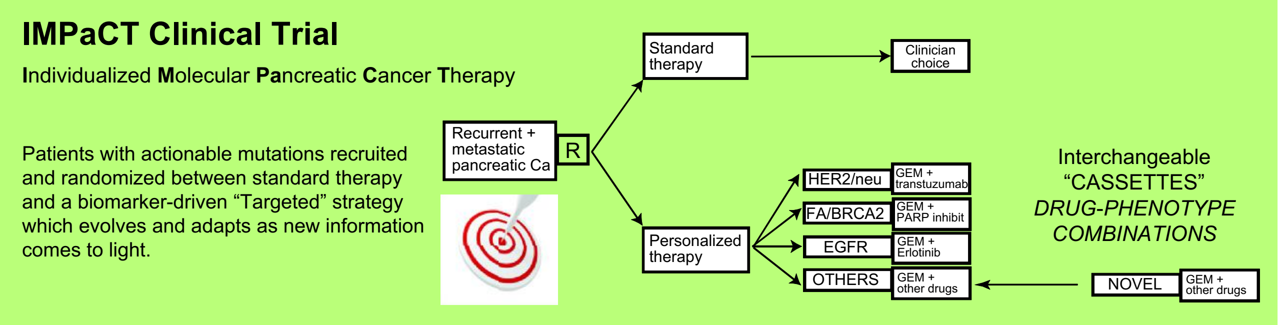

IMPaCT Clinical Trial
Individualized Molecular Pancreatic Cancer Therapy
Patients with actionable mutations recruited and randomized between standard therapy and a biomarker-driven "Targeted" strategy which evolves and adapts as new information comes to light.
Recurrent + metastatic pancreatic Ca
R
Standard therapy
Clinician choice
Personalized therapy
HER2/neu
GEM + transtuzumab
FA/BRCA2
GEM + PARP inhibit
EGFR
GEM + Erlotinib
OTHERS
GEM + other drugs
Interchangeable "CASSETTES"
DRUG-PHENOTYPE COMBINATIONS
NOVEL
GEM + other drugs

FIGURE 34.2 **IMPaCT trial.** IMPaCT is a phase II clinical trial where patients with predefined actionable molecular genotypes/phenotypes (three initially) were randomized to standard or stratified/personalized therapy. The trial was designed to be "adaptive."

metastatic setting after screening for actionable genotypes/phenotypes.

The initial three subgroups of interest were HER2 positive (*HER2/neu*-amplified), DNA-damaging agent–responsive (germline *BRCA1/2* or *PALB2* mutations), and anti-EGFR responsive (*KRAS* wild-type or *KRAS* codon 13 mutation). This trial was designed to be "adaptive" so that additional arms can be added as emerging and novel actionable molecular genotypes/phenotypes become better defined.

34.3.2 Locally Advanced Pancreatic Cancer

About 40% of patients present with nonmetastatic PC, but are unresectable. They include (1) patients with unresectable disease due to local invasion of surrounding structures, and (2) patients with borderline resectable disease, where R0 resection (microscopic clear surgical margin) is unlikely to be achieved. Evidence to define the best treatment approach for this group of patients, who have the potential for surgical resection, is lacking. The key goals are to (1) detect and recognize occult metastatic disease so that nonbeneficial surgery and other locoregional treatment modalities (e.g., radiotherapy) can be avoided, and (2) achieve R0 resection for patients who do eventually come to surgery. To achieve these goals, it is generally believed that induction chemotherapy should precede chemoradiation to prevent systemic progression and to select for nonprogressive patients who would benefit from subsequent radiotherapy. Building on this principle, the most effective chemotherapy would ideally be selected based on the actionable molecular genotypes/phenotypes discussed; other biologically relevant biomarkers that have been associated with a metastatic phenotype may be tested to tailor locoregional therapy in a trial setting (Figure 34.1, stratification point B).

Although yet to be proven, candidates include high expression of S100A2 and CXCR4 and loss of SMAD4, both of which have been associated with poor prognosis in patients with resectable disease and a distant relapse pattern rather than local recurrence [34–36]. If patients with tumors that possess a metastatic phenotype could be correctly defined, then prior knowledge that this subgroup would not otherwise benefit from the addition of radiotherapy would be beneficial. On the contrary, exploration of more aggressive surgical options could be justified in patients with truly localized disease.

34.3.3 Resectable Pancreatic Cancer

34.3.3.1 Neoadjuvant Therapy versus Surgery Upfront

Only about 10–15% of patients have "clearly resectable" disease on imaging at presentation. However, up to 30% of patients deemed "clearly resectable" after staging, suffer recurrences within six months after resection, and up to 60% succumb to disease within the first 20 months after surgery [37,38]. The pattern of failure as suggested by autopsy studies is primarily metastatic, with fewer than 30% of patients having local recurrence only [36,39]. The goals of treatment for this group are to (1) detect and recognize, or predict, occult metastatic disease so that nonbeneficial surgery and other locoregional treatment modalities (e.g., radiotherapy) can be avoided, and (2) achieve a prolonged disease-free survival after surgical resection. In order to identify patients with occult metastatic disease leading to early post-resection failure, some favor a neoadjuvant approach. Although neoadjuvant treatment in an unselected fashion may identify early progressors, it is not without associated complications. Better selection of patients for specific neoadjuvant therapies may improve overall outcomes (Figure 34.1, stratification points C and D).

Prognosis prediction tools, such as nomograms, have been developed for many cancer types in an effort to better predict prognosis and inform treatment decisions for individual patients. The most widely used tool in resected PC is the prognostic nomogram from the Memorial Sloane-Kettering Cancer Center [40], which has been tested and validated in multiple patient cohorts [41–43]. It is superior to the AJCC staging system because it segregates patients into more homogeneous prognostic groups [43], and it can be used to predict poor post-resection prognosis due to early recurrence. However, prognostic nomograms can be used only postoperatively as they include clinico-pathological variables that become available only after pathological assessment of resected specimens, by which time patients have already undergone extensive surgery of uncertain benefit.

To circumvent this, the authors' group developed a preoperative prognostic nomogram using only variables that can be measured preoperatively [44]. These include patient age, tumor size, tumor location, and tumor expression of two biologically relevant molecular biomarkers: the S100A2 and S100A4 calcium-binding proteins. The prognostic significance of these two prometastatic S100 proteins has been tested and validated in multiple patient cohorts of resectable PC ($n > 1000$), (See Chang et al., submitted for publication) and proof-of-concept data suggest that biomarker expression can be assessed in preoperative endoscopic ultrasound, fine-needle biopsy cell block samples. Our preoperative prognostic nomogram has the potential to better select patients for neoadjuvant therapy. Systemic inflammatory measures of advanced disease have also been explored, including CRP, albumin, and tumor inflammatory cell infiltrate, which may also be used to identify patients with occult metastatic disease [45–48].

34.3.3.2 Adjuvant Chemotherapy

Various large phase III RCTs have shown adjuvant chemotherapy to be beneficial [49,50], with single-agent gemcitabine having similar efficacy to that of 5-FU in unselected patients [51,52]. Questions remaining include the following: (1) Do all patients require adjuvant chemotherapy, as the prognosis of a cancer-free patient cannot be improved with chemotherapy? (2) What is the most effective chemotherapy regimen for an individual? Extrapolating from experience in the metastatic setting, there is no reason that an actionable molecular genotype/phenotype-driven adjuvant chemotherapeutic regime selection cannot be tested. The question remains as to how we can better select patients needing adjuvant chemotherapy and/or radiotherapy. The current decision for adjuvant chemotherapy is based on fitness following surgery and not on probability of residual disease. Tools such as postoperative and preoperative prognostic nomograms, described earlier, and molecular biomarkers of a metastatic phenotype can potentially be used to better select and target patients for adjuvant treatment (Figure 34.1, stratification point E).

34.3.3.3 Adjuvant Radiotherapy

Currently adjuvant radiotherapy is used sporadically, as evidence from RCTs is mixed. This is perhaps not surprising given that the majority of patients with resected PC succumb to distant metastases often with no evidence of local recurrence, which makes studies on the efficacy of radiotherapy grossly underpowered. These assertions are further supported by evidence from a meta-analysis showing that only margin-positive patients benefit from radiotherapy in addition to adjuvant chemotherapy [53].

The authors' group have demonstrated that the degree of surgical margin clearance is important for long-term survival in a subgroup of patients; this criterion can be used to identify patients who will benefit from adjuvant radiotherapy [37,38]. Molecular biomarkers of a metastatic phenotype, discussed previously, can potentially be used to select patients who are less likely to have metastatic disease and are hence more likely to benefit from more aggressive locoregional therapy.

34.4 PRECLINICAL MODELS OF PANCREATIC CANCER

Preclinical genetically engineered mouse models (GEMM) of cancer have been invaluable in contributing to our understanding of disease pathogenesis in recent years. Based on the Hingorani et al. [54] GEMM, which expresses a mutant $Kras^{G12D}$ targeted to the pancreas using Cre-Lox technology, many other models that harbor other known and novel oncogenic drivers in PC have been generated. These models closely mimic their human counterpart and can be generated to mimic actionable molecular genotypes identified in the human disease.

The GEMMs and cell lines generated are ideal resources for assessing the efficacies of predicted therapy of the putative genotype both *in vivo* and *in vitro*. In addition, the number of patient-derived xenografts (PDX) and cell lines (PDCL) that are "-omically" characterized is growing. These models have been successfully used as mouse "avatars" of the human disease that they were derived from, which can be used to test a stratified genotype-guided approach [55,56]. These preclinical models are ideal platforms to test, define, and develop novel actionable molecular genotype/phenotype and drug combinations and may serve as a basis to guide further clinical trial design and chemotherapeutic decision making when the same actionable genotype/phenotype is seen in another patient.

34.5 CENTRAL KNOWLEDGE DATABASE

With the dramatic increase in genomic sequencing facilities and the corresponding decrease in cost, the number of cancer genomes available for analysis is exponentially increasing. To tackle the complexity and heterogeneity of the cancer genome, as well as the various therapeutic regimens either on or off clinical trials, a central database of genomic sequences and treatment responses is a potential approach. As this database builds over time, evidence for the treatment of even the rarest cancer genotype may be better defined and understood. More than 100 research organizations and healthcare providers around the world have joined together to form The Global Alliance for Genomics and Health (GA4GH) to develop such a framework [57], which may potentially be incorporated into databases that match patients' molecular genotype/phenotype to all available clinical trials.

34.6 BIOREPOSITORY AND PRIVACY ISSUES

Although beyond the scope of this review, some of the greatest challenges in genotype-driven stratified medicine are those concerning privacy and biospecimen acquisition and processing. Without adequate sample acquisition, even the simplest histopathological examination cannot be performed and is often overlooked. The other highly contentious issue in the genomic era is privacy. How should the research community and healthcare providers protect patient privacy but at the same time share enough information globally to advance medical research and discovery, and in turn provide better healthcare?

34.7 CONCLUSION

In this review, we have presented an overall concept for a stratified therapeutic approach to the management of pancreatic cancer (Figure 34.1, stratification points A–E). This will continue to be an evolving area, and will need to be tested in appropriate clinical trials. It is expected that high-throughput technologies, such as massively parallel genomic sequencing, will play a key role in the future provision of stratified medicine for cancer.

REFERENCES

[1] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2013;63:11–30.

[2] Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 2012;491:399–405.

[3] Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science 2008;321:1801–6.

[4] Conroy T, Desseigne F, Ychou M, Bouche O, Guimbaud R, Becouarn Y, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Eng J Med 2011;364:1817–25.

[5] Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. N Engl J Med 2013;369:1691–703.

[6] Burris 3rd HA, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, et al. Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. J Clin Oncol 1997;15:2403–13.

[7] Young JD, Yao SY, Baldwin JM, Cass CE, Baldwin SA. The human concentrative and equilibrative nucleoside transporter families, SLC28 and SLC29. Mol Aspects Med 2013;34:529–47.

[8] Marechal R, Bachet JB, Mackey JR, Dalban C, Demetter P, Graham K, et al. Levels of gemcitabine transport and metabolism proteins predict survival times of patients treated with gemcitabine for pancreatic adenocarcinoma. Gastroenterology 2012;143:664–74. [e1–6]

[9] Farrell JJ, Elsaleh H, Garcia M, Lai R, Ammar A, Regine WF, et al. Human equilibrative nucleoside transporter 1 levels predict response to gemcitabine in patients with pancreatic cancer. Gastroenterology 2009;136:187–95.

[10] Greenhalf W, Ghaneh P, Neoptolemos JP, Palmer DH, Cox TF, Lamb RF, et al. Pancreatic cancer hENT1 expression and survival from gemcitabine in patients from the ESPAC-3 trial. J Natl Cancer Inst 2014;106:djt347.

[11] Poplin E, Wasan H, Rolfe L, Raponi M, Ikdahl T, Bondarenko I, et al. Randomized, multicenter, phase II study of CO-101 versus gemcitabine in patients with metastatic pancreatic ductal adenocarcinoma: including a prospective evaluation of the role of hENT1 in gemcitabine or CO-101 sensitivity. J Clin Oncol 2013;31:4453–61.

[12] Xia B, Dorsman JC, Ameziane N, de Vries Y, Rooimans MA, Sheng Q, et al. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. Nat Genet 2007;39:159–61.

[13] Byrski T, Huzarski T, Dent R, Gronwald J, Zuziak D, Cybulski C, et al. Response to neoadjuvant therapy with cisplatin in BRCA1-positive breast cancer patients. Breast Cancer Res Treat 2009;115:359–63.

[14] Tabernero J, Macarulla T. Changing the paradigm in conducting randomized clinical studies in advanced pancreatic cancer: an opportunity for better clinical development. J Clin Oncol 2009;27:5487–91.

[15] Ciliberto D, Botta C, Correale P, Rossi M, Caraglia M, Tassone P, et al. Role of gemcitabine-based combination therapy in the management of advanced pancreatic cancer: a meta-analysis of randomised trials. Eur J Cancer 2013;49:593–603.

[16] Gunturu KS, Yao X, Cong X, Thumar JR, Hochster HS, Stein SM, et al. FOLFIRINOX for locally advanced and metastatic pancreatic cancer: single institution retrospective review of efficacy and toxicity. Med Oncol 2013;30:361.

[17] Peddi PF, Lubner S, McWilliams R, Tan BR, Picus J, Sorscher SM, et al. Multi-institutional experience with FOLFIRINOX in pancreatic adenocarcinoma. JOP 2012;13:497–501.

[18] Clark-Knowles KV, O'Brien AM, Weberpals JI. BRCA1 as a therapeutic target in sporadic epithelial ovarian cancer. J Oncol 2010:891059.

[19] Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.

[20] Neuzillet C, Tijeras-Raballand A, Cros J, Faivre S, Hammel P, Raymond E. Stromal expression of SPARC in pancreatic adenocarcinoma. Cancer Metastasis Rev 2013;32:585–602.

[21] Infante JR, Matsubayashi H, Sato N, Tonascia J, Klein AP, Riall TA, et al. Peritumoral fibroblast SPARC expression and patient outcome with resectable pancreatic adenocarcinoma. J Clin Oncol 2007;25:319–25.

[22] Mantoni TS, Schendel RR, Rodel F, Niedobitek G, Al-Assar O, Masamune A, et al. Stromal SPARC expression and patient survival after chemoradiation for non-resectable pancreatic adenocarcinoma. Cancer Biol Ther 2008;7:1806–15.

[23] Von Hoff DD, Ramanathan RK, Borad MJ, Laheru DA, Smith LS, Wood TE, et al. Gemcitabine plus nab-paclitaxel is an active regimen in patients with advanced pancreatic cancer: a phase I/II trial. J Clin Oncol 2011;29:4548–54.

[24] Moore MJ, Goldstein D, Hamm J, Figer A, Hecht JR, Gallinger S, et al. Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. J Clin Oncol 2007;25:1960–6.

[25] da Cunha Santos G, Dhani N, Tu D, Chin K, Ludkovski O, Kamel-Reid S, et al. Molecular predictors of outcome in a phase 3 study of gemcitabine and erlotinib therapy in patients with advanced pancreatic cancer: National Cancer Institute of Canada Clinical Trials Group Study PA.3. Cancer 2010;116:5599–607.

[26] Brenton JD, Carey LA, Ahmed AA, Caldas C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? J Clin Oncol 2005;23:7350–60.

[27] Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. Lancet 2010;376:687–97.

[28] Kimura K, Sawada T, Komatsu M, Inoue M, Muguruma K, Nishihara T, et al. Antitumor effect of trastuzumab for pancreatic cancer with high HER-2 expression and enhancement of effect by combined therapy with gemcitabine. Clin Cancer Res 2006;12:4925–32.

[29] Buchler P, Reber HA, Eibl G, Roth MA, Buchler MW, Friess H, et al. Combination therapy for advanced pancreatic cancer using Herceptin plus chemotherapy. Int J Oncol 2005;27:1125–30.

[30] Chang Z, Li Z, Wang X, Kang Y, Yuan Y, Niu J, et al. Deciphering the mechanisms of tumorigenesis in human pancreatic ductal epithelial cells. Clin Cancer Res 2013;19:549–59.

[31] Safran H, Iannitti D, Ramanathan R, Schwartz JD, Steinhoff M, Nauman C, et al. Herceptin and gemcitabine for metastatic pancreatic cancers that overexpress HER-2/neu. Cancer Invest 2004;22:706–12.

[32] Harder J, Ihorst G, Heinemann V, Hofheinz R, Moehler M, Buechler P, et al. Multicentre phase II trial of trastuzumab and capecitabine in patients with HER2 overexpressing metastatic pancreatic cancer. Br J Cancer 2012;106:1033–8.

[33] Chou A, Waddell N, Cowley MJ, Gill AJ, Chang DK, Patch AM, et al. Clinical and molecular characterization of HER2 amplified-pancreatic cancer. Genome Med 2013;5:78.

[34] Biankin AV, Kench JG, Colvin EK, Segara D, Scarlett CJ, Nguyen NQ, et al. Expression of S100A2 calcium-binding protein predicts response to pancreatectomy for pancreatic cancer. Gastroenterology 2009;137:558–68.

[35] Bachet JB, Marechal R, Demetter P, Bonnetain F, Couvelard A, Svrcek M, et al. Contribution of CXCR4 and SMAD4 in predicting disease progression pattern and benefit from adjuvant chemotherapy in resected pancreatic adenocarcinoma. Ann Oncol 2012;23:2327–35.

[36] Iacobuzio-Donahue CA, Fu B, Yachida S, Luo M, Abe H, Henderson CM, et al. DPC4 gene status of the primary carcinoma correlates with patterns of failure in patients with pancreatic cancer. J Clin Oncol 2009;27:1806–13.

[37] Chang DK, Johns AL, Merrett ND, Gill AJ, Colvin EK, Scarlett CJ, et al. Margin clearance and outcome in resected pancreatic cancer. J Clin Oncol 2009;27:2855–62.

[38] Jamieson NB, Chan NI, Foulis AK, Dickson EJ, McKay CJ, Carter CR. The prognostic influence of resection margin clearance following pancreaticoduodenectomy for pancreatic ductal adenocarcinoma. J Gastrointest Surg 2013;17:511–21.

[39] Barugola G, Falconi M, Bettini R, Boninsegna L, Casarotto A, Salvia R, et al. The determinant factors of recurrence following resection for ductal pancreatic cancer. JOP 2007;8:132–40.

[40] Brennan MF, Kattan MW, Klimstra D, Conlon K. Prognostic nomogram for patients undergoing resection for adenocarcinoma of the pancreas. Ann Surg 2004;240:293–8.

[41] de Castro SM, Biere SS, Lagarde SM, Busch OR, van Gulik TM, Gouma DJ. Validation of a nomogram for predicting survival after resection for adenocarcinoma of the pancreas. Br J Surg 2009;96:417–23.

[42] White RR, Kattan MW, Haney JC, Clary BM, Pappas TN, Tyler DS, et al. Evaluation of preoperative therapy for pancreatic cancer using a prognostic nomogram. Ann Surg Oncol 2006;13:1485–92.

[43] Ferrone CR, Kattan MW, Tomlinson JS, Thayer SP, Brennan MF, Warshaw AL. Validation of a postresection pancreatic adenocarcinoma nomogram for disease-specific survival. J Clin Oncol 2005;23:7529–35.

[44] Chang D, Colvin E, Scarlett C, Pajic M, Pinese M, Gill A. A molecular prognostic nomogram for resectable pancreatic cancer. J Clin Oncol 2011;29:154.

[45] Jamieson NB, Mohamed M, Oien KA, Foulis AK, Dickson EJ, Imrie CW, et al. The relationship between tumor inflammatory cell infiltrate and outcome in patients with pancreatic ductal adenocarcinoma. Ann Surg Oncol 2012;19:3581–90.

[46] Jamieson NB, Denley SM, Logue J, MacKenzie DJ, Foulis AK, Dickson EJ, et al. A prospective comparison of the prognostic value of tumor- and patient-related factors in patients undergoing potentially curative surgery for pancreatic ductal adenocarcinoma. Ann Surg Oncol 2011;18:2318–28.

[47] Glen P, Jamieson NB, McMillan DC, Carter R, Imrie CW, McKay CJ. Evaluation of an inflammation-based prognostic score in patients with inoperable pancreatic cancer. Pancreatology 2006;6:450–3.

[48] Jamieson NB, Glen P, McMillan DC, McKay CJ, Foulis AK, Carter R, et al. Systemic inflammatory response predicts outcome in patients undergoing resection for ductal adenocarcinoma head of pancreas. Br J Cancer 2005;92:21–3.

[49] Neoptolemos JP, Stocken DD, Friess H, Bassi C, Dunn JA, Hickey H, et al. A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. N Engl J Med 2004;350:1200–10.

[50] Oettle H, Post S, Neuhaus P, Gellert K, Langrehr J, Ridwelski K, et al. Adjuvant chemotherapy with gemcitabine vs observation in patients undergoing curative-intent resection of pancreatic cancer: a randomized controlled trial. JAMA 2007;297:267–77.

[51] Neoptolemos JP, Stocken DD, Bassi C, Ghaneh P, Cunningham D, Goldstein D, et al. Adjuvant chemotherapy with fluorouracil plus folinic acid vs gemcitabine following pancreatic cancer resection: a randomized controlled trial. JAMA 2010;304:1073–81.

[52] Regine WF, Winter KA, Abrams RA, Safran H, Hoffman JP, Konski A, et al. Fluorouracil vs gemcitabine chemotherapy before and after fluorouracil-based chemoradiation following resection of pancreatic adenocarcinoma: a randomized controlled trial. JAMA 2008;299:1019–26.

[53] Stocken DD, Buchler MW, Dervenis C, Bassi C, Jeekel H, Klinkenbijl JH, et al. Meta-analysis of randomised adjuvant therapy trials for pancreatic cancer. Br J Cancer 2005;92:1372–81.

[54] Hingorani SR, Wang L, Multani AS, Combs C, Deramaudt TB, Hruban RH, et al. Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. Cancer Cell 2005;7:469–83.

[55] Dennis C. Mouse ’avatars’ could aid pancreatic cancer therapy. Nat News 2012. http://dx.doi.org/10.1038/nature.2012.10259.

[56] Hidalgo M, Von Hoff DD. Translational therapeutic opportunities in ductal adenocarcinoma of the pancreas. Clin Cancer Res 2012;18:4249–56.

[57] The Global Alliance for Genomics and Health (GA4GH). <http://oicr.on.ca/oicr-programs-and-platforms/global-alliance-genomics-and-health-ga4gh>. [accessed 03.01.14].

Chapter 35

Pharmacogenomics in Anesthesia

Shiv Kumar Singh[a] **and Tushar Dixit**[b]

[a]*Royal Liverpool University Hospitals, UK,* [b]*Mersey Deanery, UK*

Chapter Outline

35.1 Overview 815
35.2 Pharmacokinetic Pharmacogenetics In Anesthesia 815
35.2.1 Transport Proteins 815
35.2.2 Metabolism 816
35.3 Pharmacodynamics 820
35.3.1 Defective Enzymes and Anesthesia 820
35.3.2 Abnormalities in Butyrylcholinesterase 821
35.3.3 Malignant Hyperthermia 822
35.3.4 Pharmacogenetics of Specific Drugs in Anesthesia 825
35.3.5 Inhalational Anesthetic Agents 827
35.3.6 Neuromuscular Blocking Agents 828
35.3.7 Opioids 828
35.3.8 Nonsteroidal Anti-Inflammatory Drugs 830
35.4 Conclusion 830
References 830

35.1 OVERVIEW

Rare pharmacogenetic adverse effects—for example, induced apnea, malignant hyperthermia, and thiopentone-induced acute porphyria—are well recognized. As early as the 1950s, Kalow and Gunn described how an inherited variation in drug metabolism involving the enzyme butyrylcholinesterase affects the response to succinylcholine [1]. However, expanding pharmacogenetic knowledge of variable efficacy and adverse effects related to other commonly prescribed agents (e.g., analgesics, antiemetics, beta-blockers) will have a major impact on anesthetic practice in the future [2–4]. In February 2013, for example, the FDA issued a box warning against codeine for pain relief following tonsillectomies in children. Using the database of the Adverse Events Reporting System (AERS) from 1969 to 2012, the FDA identified ten deaths and three overdoses associated with Codeine [5]

In this chapter, we describe how genetic variations have a major impact on efficacy and toxicity of drugs commonly used in anesthesia.

35.2 PHARMACOKINETIC PHARMACOGENETICS IN ANESTHESIA

Many important examples in pharmacogenetics relate to pharmacokinetic parameters associated with variable drug uptake, metabolism, or elimination (see Chapter 17).

35.2.1 Transport Proteins

Transmembrane proteins carry either endogenous compounds or xenobiotics across biological membranes. They can be classified as either efflux or uptake proteins, depending on the direction of transport. The extent of expression of genes coding for transmembrane proteins can have a profound effect on the bioavailability and pharmacokinetics of various drugs. Transmembrane proteins, particularly of the gut, liver, kidney, and blood–brain barrier, have been implicated in multidrug resistance in epilepsy and cancer chemotherapy.

There are two superfamilies of transport proteins that have important effects on the absorption, distribution, and excretion of drugs. These are the ATP-binding cassette

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00035-9
© 2014 Elsevier Inc. All rights reserved.

(ABC) and the solute-carriers (SLCs). ABC transporters are present in cellular and intracellular membranes and can be responsible for either importing or removing (efflux) substances from cells and tissues. They often transport substances against a concentration gradient using energy from the hydrolysis of ATP to drive the transport. There are at least 49 ABC transporter genes, of which ABCB1 is of importance to anesthesia. ABCB1 codes for a glycosylated membrane protein commonly referred to as P-glycoprotein (P-gp), PGY1, or multidrug resistance protein-1 (MDR1).

35.2.1.1 P-glycoprotein and Anesthesia

ABCB1/MDR1 was the first recognized and most extensively studied ABC transporter encoding a polypeptide (P-glycoprotein). P-glycoprotein (P-gp) is one of the main efflux transporters (pump) proteins and is expressed in multiple normal tissues with excretory or protective functions, including intestine, kidney, liver, blood–brain barrier, spinal cord, testes, and placenta [6,7]. The expression of ABCB1 on the apical surface of enterocytes and the canalicular part of the hepatocyte has been shown to influence intestinal drug absorption and limit oral bioavailability of a wide variety of structurally diverse drugs.

ABCB1 appears to have a superior effect on decreasing drug cellular uptake from the blood into the brain as well as from the intestinal lumen into the epithelial cells, as opposed to increasing drug elimination from hepatocytes and renal tubules into adjacent luminal spaces. ABCB1 is highly polymorphic, with more than 50 SNPs and three insertion/deletion polymorphisms reported [8–10]. There is a broad overlap in substrate specificity and tissue distribution of P-gp with cytochrome P450 3A4 (CYP3A4), as discussed later in the chapter.

This broad specificity, coupled with the tissue localization and function of both proteins, has led to the hypothesis that P-gp and P450 3A4 work in concert, protecting the body from absorption of harmful compounds by acting synergistically in the small intestine. There is considerable interindividual variability (2–8-fold) in human P-gp expression, and numerous polymorphisms and haplotypes in humans have been identified [11]. The frequencies of these SNPs in a population can vary according to racial/ethnic background and gender. For instance, hepatic P-gp levels are 2–2.4-fold lower in females than in males. Thus, females may have not only increased response but also increased toxicity.

Some opioid analgesics, such as morphine, methadone, sufentanil, and fentanyl, are substrates for P-gp. Absence or blockade of this protein can lead to an increased response to substrate drugs [12]. Clinically, simultaneous administration of P-gp inhibitors and oral morphine can increase the plasma concentration and analgesic effect of morphine. Investigation into MDR-1 polymorphisms may promote understanding of variability in the effect of drugs that are P-gp substrates [13]. Various cancers tend to display low initial P-gp, with expression increasing after chemotherapy and relapse. This may explain why these patients need higher doses of oral opioids and suffer from the side effects like constipation, nausea, and vomiting.

In their preoperative assessment of patients, anesthetists may come across many drugs that the patient may be on for various illnesses that are P-gp substrates. These include protease inhibitors, immunosuppressants, calcium channel blockers, beta-blockers, statins, steroids, antihistamines, anticonvulsants, and antidepressants. The importance of P-gp for pharmacotherapy has led to great interest in its pharmacogenomics.

35.2.2 Metabolism

35.2.2.1 Phase I Enzymes

Many phase I reactions, particularly oxidative pathways, occur in the liver because of a nonspecific mixed-function oxidase system in the endoplasmic reticulum. This is the cytochrome P450 system (CYP), named after the wavelength (in nm) of its enzymes' maximal absorption of light when the reduced state is combined with carbon monoxide. However, the cytochrome P450 system is not unique to the liver; its enzymes are also found in gut mucosa, lung, brain, and kidney. CYPs are responsible for 70–80% of phase 1 metabolism of medications [14].

CYPs are a superfamily of drug-metabolizing enzymes and are classified into families and subfamilies by their degree of shared amino acid sequences. Fifty-seven CYP genes have been identified in humans, but only a relatively small number of the encoded proteins, primarily the CYP1, CYP2, and CYP3 families, appear to contribute to drug metabolism. The clinical relevance of the inherited variations associated with CYP enzymes is dependent on many environmental factors, such as disease, drugs, surgery, nutritional status, and biological variations. The frequency of variant alleles is also significantly varied depending on race and ethnic background [15].

CYPs that share >40% amino acid sequence are grouped in a family, which is designated by a number (e.g., CYP2); those with >55% homology are grouped in a subfamily, which is designated by a letter (e.g., CYP2A). In simple terms, families and subfamilies respectively share 40% and 55% of the amino acid sequence. Subfamilies are further divided into isoforms, which are identified by a third number (e.g., CYP2A6). Allelic variants are designated by an asterisk (*) and a number following the protein identifier (e.g., CYP2A6*1). The majority of drugs are metabolized by CYPs 1, 2, and 3—mainly CYPs 2C, 2D6, and 3A. CYP2C9 and CYP2D6 both demonstrate significant genetic polymorphism; other cytochromes have variants as well, but these two are clinically important. Many drugs are metabolized by more than one isozyme (e.g., diazepam by CYP3A4 and CYP3A5) [14,16,17].

CYP2EI

CYP2E1 is responsible for the metabolism of inhalational agents and paracetamol [3,18]. Variability has not been associated with genetic polymorphism as such but rather with body mass, diet, alcohol consumption, and age.

CYP3A

The CYP3A family of CYPs is responsible for the oxidative metabolism of more than 50% of all therapeutic medications that undergo phase I metabolism. In adults this enzyme comprises CYP3A4 and CYP3A5. CYP3A4 is found in large quantities in the epithelial layer of the small intestine and in the liver (~30% of all CYPs).

CYP3A4 has the ability to metabolize a variety of drugs from many different classes: opioids, benzodiazepines, local anesthetics, immunosuppressants, calcium channel anatagonists, and aantihistamines [2,3]. This enzyme exhibits wide variability in activity and continuous unimodal distribution. Although dose adjustments are generally not needed for CYP3A4 polymorphisms, drugs metabolized through CYP3A4 are subject to inhibitors such as cimetidine, ketoconazole, and grapefruit juice. This inhibition leads to reduced CYP3A4 activity.

CYP3A5

CYP3A5 is similar to CYP3A4 and metabolizes similar substrates, but its activity is diminished. Recent findings suggest that hepatic (and intestinal) CYP3A5 expression may be much higher than what was once believed (>50% of hepatic CYP3A). Individuals expressing CYP3A5 along with 3A4 may have greater metabolism of CYP3A drugs.

CYP2D

CYP2D is another important member of the CYP family of which two isoforms, CYP2D6 and 2D9, are of importance to anesthesia. Even though it accounts for only 2–5% of all the CYPs, CYP2D6 metabolizes ~25% of phase I drug reactions. This enzyme is noted for extensive variability and for its effects on the metabolism of many anesthetic drugs. There may be as much as a 1000-fold difference in metabolism by CYP2D6 between phenotypes, which may result in adverse drug reactions in patients with a polymorphism of the CYP2D6 enzyme. The CYP2D6 enzyme is responsible for metabolism of many antiemetics (5HT-3 antagonists), beta-blockers, opioids that undergo CYP2D6-mediated O-demethylation (codeine, dihydrocodeine, tramadol, oxycodone, and hydroxycodone), antipsychotics, anti depressants (SSRIs), neuroleptics, and antiarrhythmics [2–4,19–22].

Polymorphic variation in CYP2D6 can result in patients being classified as poor (PM), intermediate (IM), extensive (EM, wild-type), and ultrarapid (UM) metabolizers. Although testing may be beneficial in anesthesia, associated adjustments in dosing and clear recommendations have not been fully developed and have not become standard practice. Patients with poor metabolism have nonfunctional or nonexistent CYP2D6 enzymes, and those with ultrarapid metabolism have overactive enzymes, requiring a modification of drug dose depending on whether the drug is activated or metabolized by the enzyme [22].

The overall frequency of PM is 7–8% in Caucasians, 2% in African Americans, and <1% in Asians. More than 20 mutant CYP2D6 alleles have been identified. Most code for inactive proteins, and homozygotes are PMs. However, some alleles that are particularly frequent in Asians code for a protein with reduced activity. Ultrarapid metabolism is caused by gene duplication, with up to 13 copies. The frequency of UMs is 20% in Ethiopians, 7% in Spanish, and 1% in Scandinavians, generally requiring substantially higher drug doses. There is even a spectrum of CYP2D6 activity in EMs owing to the diversity of 2D6 alleles [4,22–24].

The role of pharmacogenetics in interindividual variability in drug disposition and response is well illustrated by codeine, a prodrug requiring CYP2D6-mediated metabolic activation (O-demethylation) to the μ (mu) agonist morphine, and even greater bioactivation to the more potent agonist morphine-6-glucuronide. Patients who rapidly convert codeine into morphine usually have at least one wild-type CYP2D6 allele, known as CYP2D6*1. In contrast, poor codeine metabolizers tend to express combinations of the *4, *5, and *6 alleles and are virtually immune to codeine's analgesic effects [25]. Interestingly, the severity of the adverse effects of codeine does not seem to be associated with the CYP2D6 genotype, thus explaining how codeine can impair driving ability even when the driver's blood does not contain detectable morphine concentrations [26].

CYP2D9

Drugs metabolized by CYP2D9 include pheyntoin, NSAIDS, celecoxib, and warfarin. Testing is available for this enzyme, and several studies have shown that performing and evaluating CYP2D9 genotyping can be helpful in providing safe, more personalized warfarin treatment and in reaching therapeutic levels sooner, at the same time reducing potentially life-threatening side effects [2,27].

CYP2C

The CYP2C gene family consists of CYPs 2C8, 2C9, 2C18, and 2C19, which together account for ~25% of total hepatic CYP and metabolize 15–20% of all therapeutic drugs. CYP2C9 is the major 2C isoform in human liver. The best known CYP2C9 substrates are (S)-warfarin, phenytoin, tolbutamide, diclofenac, and several NSAIDs [3,20,28,29]. Several variant 2C9 proteins with substantially decreased activity (as low as 20%) have been identified. These mutant alleles occur more commonly in Caucasians (10–20%) than in Asians and African Americans (<5%). Warfarin is an excellent example of pharmacogenetics and its application in modern therapeutics. Patients with CYP2C9 variants have higher warfarin plasma concentrations and are more susceptible to over-anticoagulation, serious or life-threatening bleeding complications, and longer hospitalizations [27].

CYP2C19

CYP2C19 is a minor CYP, but it has been well studied. Its substrates include barbiturates, diazepam, phenytoin, amitryptyline, and proton pump inhibitors. Individuals are classified as EM or PM, with a PM frequency of 2–5% in Caucasians, 2% in African Americans, 18–23% in Japanese, and 15–17% in Chinese. The CYP2C19 polymorphism is important in the disposition and clinical effect of diazepam, particularly in PM homozygotes. PM diazepam clearance is half that in EMs, with dose requirements correspondingly less; heterozygotes are IMs, which is thought to account for lower diazepam dose requirements in Chinese patients [17].

35.2.2.2 Modified Drug Metabolism

The hepatic CYP system is particularly subject to enzyme induction. In other words, exposure to environmental substances or drugs can increase expression of the gene coding for the enzyme and thus the amount of the enzyme present. This is one of the most important mechanisms of drug interaction.

Induction may be of one or several of the CYP subtypes, and it may increase metabolism of the inducing substance or of another drug, thus increasing its excretion and reducing its effect. There are several examples of the relevance of induction to anesthetic practice. The toxicity of pethidine may be enhanced by barbiturates, increasing its metabolism to the toxic metabolite norpethidine. Monoamine oxidase inhibition (MAO-I) and alanine (Ala) synthase induction in porphyria are other examples.

Inhibition of metabolism may also occur that may be due to inhibition of enzymes or, more commonly, to competition between drugs for the same metabolic pathway. With most drugs that are eliminated by processes obeying first-order kinetics, competition rarely causes significant metabolism inhibition, but competition is more important in drugs eliminated by zero-order kinetics, where the enzyme system is easily saturated. For example, omeprazole can inhibit the metabolism of drugs metabolized mainly by the cytochrome P-450 enzyme subfamily 2C (diazepam, phenytoin), but not that of drugs metabolized by subfamilies 1A (caffeine, theophylline), 2D (metoprolol, propranolol), and 3A (ciclosporin, lidocaine (lignocaine), quinidine). Since relatively few drugs are metabolized mainly by the 2C compared with the 2D and 3A families, the potential for omeprazole to interfere with the metabolism of other drugs appears to be limited, but the half-life of diazepam and phenytoin is prolonged as much as by cimetidine [30,31].

35.2.2.3 Non-CYP Enzymes

The CYP system is not responsible for all of phase I metabolism. Many non-CYP enzymes are involved in the metabolism of drugs used in anesthesia.

Catechol-O-methyltransferase

The monoamines (adrenaline, noradrenaline, dopamine) are metabolized by the mitochondrial enzyme monoamine oxidase. Individual genetic variation, or the presence of exogenous inhibitors of this breakdown pathway, can result in high levels of monoamines in the circulation, causing severe cardiovascular effects. Catechol-O-methyltransferase (COMT) regulates the metabolism of catecholamines such as norepinephrine, epinephrine, and dopamine; thus, mutations in the COMT gene can effect the perception of pain. Reduced COMT activity leads to higher levels of catecholamine neurotransmitters and increased pain sensitivity in animal models of pain [32]. A common human polymorphism in the COMT gene, p.V158M (Val158Met), leads to a 3–4-fold variation in COMT activity, with p.158MM homozygous individuals demonstrating the lowest activity and subsequent reduced activation of the endogenous opioid system and compensatory up-regulation of opioid receptor expression [33]. Although low COMT activity is associated with increased pain sensitivity in some patients, clinical studies of COMT polymorphisms suggest that these individuals require smaller opioid doses, possibly because of up-regulation of opioid receptors. Among the cancer population, carriers of the COMT, p.158VV, and p.158VM genotype have been shown to require 63% and 23% higher doses of morphine compared with carriers of the p.158MM genotype. In the postoperative pain setting, Kim and colleagues failed to show a significant association between COMT polymorphisms and pain response, although a low frequency of some polymorphisms and lack of statistical power may have contributed to this result [34].

Carboxylesterases

Carboxylesterases (CESs) have a wide tissue distribution and are found in the greatest amounts in the liver and in the gastrointestinal tract, brain, and possibly blood. CESs belongs to the class of serine hydrolases that include acetylcholinesterase, which is primarily found in the blood and neural synapses. They are categorized as phase I drug-metabolizing enzymes that can hydrolyze a variety of ester-containing drugs and prodrugs, such as angiotensin-converting enzyme inhibitors (ACE-I) and opioids (cocaine, heroin, and pethidine). It is thought that CESs are one of the major determinants for pharmacokinetics and pharmacodynamics of ester drugs or ester prodrugs. There are numerous polymorphisms in the gene, but none result in altered enzyme activity [35].

Butyrylcholinesterase

A major human cholinesterase, butyrylcholinesterase (also known as pseudocholinesterase, plasma cholinesterase, serum cholinesterase, and nonspecific cholinesterase), is found in plasma as well as liver and other tissues. Plasma cholinesterase was one of the early enzymes for which pharmacogenetic variation was elucidated based on observed heritable variability in response to succinylcholine. (discussed in detail later in the chapter). The level and quality of butyrylcholinesterase (BChE) in a patient determines the duration of action of succinylcholine and mivacurium. Inherited deficiency or reduced efficacy of plasma cholinesterase results in prolonged muscle relaxation after succinylcholine [36]. This was the first documented example of inherited variation in anesthetic drug effect [1,37–40]. Plasma cholinesterase also metabolizes the three ester local anesthetics: cocaine, procaine, and chloroprocaine. Heroin undergoes extensive deacetylation in blood to 6-monoacetylmorphine and then to morphine. The former reaction is catalyzed rapidly by plasma cholinesterase and more slowly by erythrocyte AchE, although the latter reaction accounts for most hydrolysis *in vivo* and only erythrocyte AchE can convert 6-monoacetylmorphine to morphine [1,41].

Like succinylcholine, esmolol and remifentanil also undergo extensive metabolic inactivation in blood; their metabolism is catalyzed exclusively by erythrocyte AchE. Plasma cholinesterase, which catalyzes succinylcholine, has no role to play; thus, neither remifentanil nor esmolol elimination is altered by pseudocholinesterase deficiency [13].

35.2.2.4 Phase II Enzymes

Phase II reactions are conjugation reactions in which the drug or its metabolite is attached to a water-soluble molecule or group, such as glucuronic acid, glutathione, sulfate, acetyl group, methyl group, or glucosamine, making the whole complex more hydrophilic for excretion via the renal system. The role of phase II enzyme polymorphisms in the metabolism of anesthetic drugs is largely unknown, but some of the important drugs administered to patients in the perioperative period by anesthetists undergo phase II reactions.

N-acetylransferases

N-acetyltransferases (NAT1 and NAT2) are key enzymes in the conjugation (phase II reaction) of certain drugs and other xenobiotics with an arylamine structure [42]. These enzymes have historical significance as they were the first drug-metabolizing enzymes for which genetic polymorphisms were discovered (fast and slow acetylators for isoniazid). Slow acetylators show a greater therapeutic response than fast acetylators to several drugs (e.g., isoniazid, hydralazine, dapsone) and may also be more susceptible to the side effects mediated by acetylated metabolites, such as isoniazid hepatotoxicity and lupus-like syndrome after hydralazine or procainamide. Fast acetylators, on the other hand, may not reach therapeutic levels with treatment and require greater doses [43].

Polymorphisms in NAT2 have long been recognized to modulate toxicity produced by the antitubercular drug isoniazid, with molecular epidemiologic studies suggesting a link between the acetylator phenotype and increased risk for bladder cancer. Polymorphisms in the NAT2 gene are responsible for the N-acetylation polymorphism that divides human populations into rapid, intermediate, and slow acetylator phenotypes. Multiple NAT2 alleles (NAT2*5, *6, *7, and *14) have substantially decreased acetylation activity and are common in Caucasians and populations of African descent. Polymorphisms in NAT1 are generally associated with relatively minor effects on acetylation function and less interindividual variability than seen with NAT2 polymorphisms [44,45].

UDP-glucuronosyltransferases and Glutathione S-Transferases

UDP-glucuronosyltransferases (UGTs) and glutathione S-transferases (GSTs) are present predominantly in the gastrointestinal tract, especially in the liver; however, they have also been distinguished in extrahepatic tissues, including skin, kidney, intestine, and many other organs. UGTs and GSTs are important families of detoxification enzymes. UGTs catalyze the conjugation of a wide variety of exogenous compounds (e.g., drugs, pesticides, components of tobacco smoke) and endogenous compounds (e.g., bilirubin, bile acids, steroid hormones) to glucuronic acid, while GSTs catalyze the reaction of glutathione with exogenous electrophiles (e.g., polycyclic aromatic hydrocarbons, heterocyclic amines) and endogenous products of oxidative stress. The metabolites formed by these reactions are generally less toxic and more water-soluble, which facilitates their biliary and renal excretion. Polymorphic variations in these detoxification enzymes may influence the rates of conversion of toxic or carcinogenic compounds [46,47].

In humans, UGT have been mainly classified into two families: UGT1A and UGT2. So far, nine functional UGT1A isoenzymes (UGT1A1, UGT1A3–UGT1A10) have been characterized, all deriving from a single gene locus on chromosome 2. UGT1A enzymes are involved mainly in the metabolism of exogenous compounds; although this is not strictly the case, as bilirubin and steroid hormones are important endogenous substrates. UGT2 isoenzymes 2A1, 2B4, 2B7, 2B10, 2B11, 2B15, and 2B17 are involved mainly in the glucuronidation of endogenous compounds. Several functional polymorphisms in the UGT1A family of isoenzymes have been identified. In Caucasians, polymorphisms in this family result in significantly reduced hepatic bilirubin UGT enzyme activity, leading to a mild form of hyperbilirubinemia known as Gilbert's syndrome [48,49].

Glucuronidation is an important pathway for certain drugs used in anesthesia. Propofol is mainly eliminated via glucuronidation by human UGT1A9. Many opioids such as morphine, codeine, naloxone, nalorphine, buprenorphine, oxymorphone, and hydromorphone undergo glucuronidation by UGT2B7. Glucuronidation of morphine to morphine 6-glucuronide (M6G) is important since this metabolite is more potent than its parent drug. The finding that UGT2B7 is also present in human brain suggests that in situ formation of M6G may play an important role in analgesia [50]. Midazolam is metabolized to 1-hydroxymidazolam by glucuronidation and is pharmacologically active. It is excreted renally and may be responsible for the prolonged effects of midazolam in patients with renal insufficiency when used for ITU sedation [3].

In humans, the GST family comprises four main classes (alpha, mu, pi, theta), whose genes are mapped on different chromosomes. These enzymes catalyze the reaction of the tripeptide glutathione with a diverse array of drugs and toxins, and they are primarily a defensive system for detoxification. There is considerable polymorphism in GST expression; however, this enzyme is of little relevance to anesthesia (except in the metabolism of the sevoflurane degradation product “compound A” [3,49].

35.3 PHARMACODYNAMICS

Pharmacodynamic polymorphisms are generally attributable to variants in cell-surface receptors or intracellular signaling pathways Compared with genetic influences on drug disposition, less information is currently available about genetic influences on pharmacodynamics. There are, however, several notable examples. Malignant hyperthermia is due to polymorphism in the skeletal-muscle calcium release channel (ryanodine receptor) gene.

Anesthetists are well aware of the interindividual variability of response to morphine, and it is of considerable interest that this may be due to opioid receptor polymorphisms in a certain group of patients [50]. Among the numerous candidate genes that have been considered important in opioid response, the μ-opioid receptor gene (OPRM1, p.118A/G), the catechol-O-methyltransferase gene (COMT, Val158Met), several variants of the ATP-binding cassette, subfamily B member 1 gene (ABCB1), and the CYP family of enzymes have been extensively reviewed.

The level of evidence linking gene variation to morphine response is moderate. In patients with the A118G μ opioid receptor polymorphism, some but not all studies show decreased A118G μ receptor affinity for ß-endorphin and morphine. Some have shown a 2–3-fold decrease in the potency of morphine and morphine-6-glucuronide, while others have shown no difference. Some have shown a 1.5–2-fold increase in postoperative morphine PCA requirements in A118G homozygotes (but not heterozygotes), while others suggest no major difference [32,33,51–55].

In the past decade, the use of β-blockers has become a standard of care in at-risk patients, with the goal of improving perioperative outcomes. However, despite numerous *in vitro* studies that have extensively defined the biological effects of the multiple genetic variants of the β1-adrenergic receptor gene (ADRB1) and the β2-adrenergic receptor gene (ADRB2), the clinical relevance of most of the numerous pharmacogenetic results for β-blockers is still inconclusive [2,56].

Several nonsynonymous polymorphisms in adrenergic receptors have been identified, although their functional significance awaits full determination. Some ß-adrenergic receptor polymorphisms have been implicated in the response to asthma therapy, and it is speculated that they are involved in the response to drugs used for hypertension and congestive heart failure [2].

35.3.1 Defective Enzymes and Anesthesia

A different type of problem arises when there are genetically determined defects in enzyme activity, as in porphyria, or defects at the receptor or intracellular level as in malignant hyperpyrexia. A detailed description of these conditions can be found in any standard anesthesia textbook or in review articles.

35.3.1.1 Porphyria

Porphyria comprises a group of metabolic disorders resulting from partial enzyme deficiency in the haem-biosynthesis pathway that are characterized by acute neurovisceral symptoms, skin lesions, or both. Acute porphyrias (acute intermittent porphyria, variegate porphyria, or hereditary coproporphyria) can be precipitated by drugs used in anesthesia and present with acute attacks, typically consisting of severe abdominal pain, nausea, constipation, confusion, and seizure. They can be life-threatening [57].

The control of haem production is tightly regulated through feedback inhibition of ALA synthesis so that the levels of haem closely match demand. The implications of this for the anesthetist is that the disease may be dormant until periods of stress occur, caused by surgery and illness or by the administration of drugs, in particular barbiturates (thiopentone). Acute intermittent porphyria (AIP) is caused by a deficiency in porphobilinogen PBG deaminase. It is an autosomal dominant disease with the gene loci on chromosome 11. AIP is of primary importance to the anesthetist because it is the porphyria in which an acute attack is most likely to prove fatal. Variegate porphyria is due to a deficiency in protopophyrinogen oxidase. The gene for this enzyme is found on chromosome 1. The genes encoding the deficient enzymes in hereditary and plumboporphyria are both situated on chromosome 9 [57,58].

Anesthetic Drugs and Porphyria

Barbiturates are the classic example of drugs that are dangerous in patients with porphyria. Propofol is the best choice at present for intravenous induction, although there is some doubt about its use as an infusion. From long use, halothane has proven safe, and isoflurane is probably so; enflurane should be avoided, and information on sevoflurane and desflurane is insufficient. Opioids, such as morphine and fentanyl, are safe, whereas there is not sufficient data on some other analgesics to be sure of their position. All muscle relaxants are probably safe, although, again, there are not enough data about most to be completely sure; on the other hand, atropine and neostigmine are known to be safe. Drugs that are unsafe or probably so include barbiturates, etomidate, enflurane, alcuronium, mepivacaine, pentazocine, some benzodiazepines (temazepam is safe; others are less certain), calcium channel blockers, and aminophylline [59,60].

In the event of an acute crisis, it is vital to assess muscle strength and bulbar function because severe disease may cause respiratory failure and risk of aspiration. Postoperative ventilation may be necessary. All possible precipitants must be withdrawn and adequate hydration must be ensured. Haematin is the only specific treatment that probably has an effect by supplementing the free haem pool and suppressing ALA synthetase. However, it may cause renal failure, coagulopathy, and thrombophlebitis. Haem arginate is a suitable alternative that is devoid of the side effects associated with haematin [61]. β-blockers have a dual role, both in controlling tachycardia and in decreasing ALA synthase activity. In the ICU setting, it is recommended that high carbohydrate intake be instituted because carbohydrate loading can suppress porphyrin synthesis [62].

Diagnosis is essential to enable specific treatments to be started as soon as possible. Screening of families to identify presymptomatic carriers is crucial to decrease the risk of overt acute porphyrias through counseling on avoidance of potential precipitants [62].

35.3.2 Abnormalities in Butyrylcholinesterase

Succinylcholine-induced neuromuscular blockade can be significantly prolonged if the patient has an abnormal genetic variant of butyrylcholinesterase (BChe). Kalow and Genest found this variant to respond to dibucaine differently from how normal butyrylcholinesterase responds. Dibucaine inhibits normal butyrylcholinesterase to a far greater extent than it inhibits the abnormal enzyme. This observation led to development of the test for dibucaine number [38]. The prevalence of BChE deficiency is highest in the Caucasian population, with 3.4–4% displaying a partial enzyme deficiency leading to slightly prolonged apnea (between 5 minutes and 1 hour) and 1 in 2500 individuals showing a prolongation of more than 1 hour. Individuals with undetectable levels of BChE activity display a severe prolongation lasting more than 8 hours. The prevalence of this severe form is estimated at 1 in 100,000 individuals [37,63].

The molecular biology of butyrylcholinesterase is well understood. Heterozygous (single-allele) expression is responsible for all plasma cholinesterase variants. The plasma cholinesterase gene lies on the long (q) arm of chromosome 3. The amino acid sequence of the enzyme is known, and the coding errors responsible for most genetic variations have been identified (more than 50 mutations). Most variants are due to a single amino acid substitution error or sequencing error at or near the active site of the enzyme [40].

35.3.2.1 Quantitative Variants

The K variant is caused by a point mutation resulting in the substitution of alanine by threonine at position 539 in the amino acid sequence of the subunit. It is the most common clinically significant BChE variant. It was recently shown that patients with a K variant have a significantly prolonged duration of mivacurium action—30–40% if they are of the U/K genotype. The J variant is associated with a 33% reduced BChE concentration.

35.3.2.2 Qualitative Variants

The A variant is a result of a point mutation resulting in aspartamine being substituted by glycine at position 70. Patients who are homozygous for the A variant are very sensitive to mivacurium and suxamethoium, and an extremely long duration of action has been reported after 0.12–0.2 mg/kg—the time for full spontaneous recovery of neuromuscular function is 6–8 hours compared to 30 minutes in patients with normal BChE. In the case of suxamethonium, the effect may persist for more than 2 hours. The F allele has reduced activity *in vivo* and demonstrates fluoride inhibition.

The S variant results from a stop codon. It occurs because of multiple mutations and has no activity. Neuromuscular blockade may last for several hours in the homozygote.

Butyrylcholinesterase activity is expressed as dibucaine number. The activity of atypical cholinesterase is decreased by 70%. Homozygotes are very sensitive to succinylcholine, and the frequency of this variant is about 2%. A small number of patients (0.3%), however, express a fluoride-resistant enzyme whose activity is decreased by 60% and who are moderately sensitive to succinylcholine. Patients with the silent variant (0.03%) have no enzyme activity. Succinylcholine hydrolysis is rapid (90% within one minute) and robust (70% reduction in enzyme activity only moderately prolongs neuromuscular blockade) [37,38,40,63]. (See Figure 35.1.)

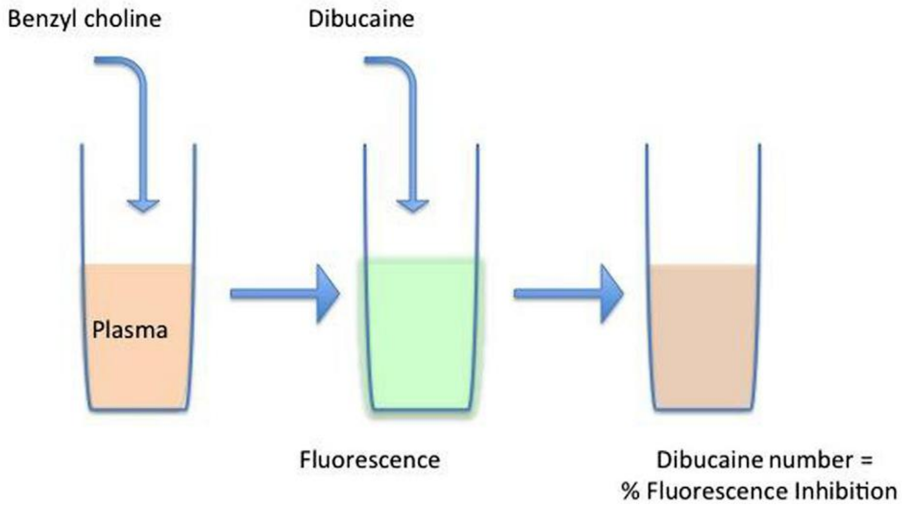


FIGURE 35.1 Dibucaine inhibits normal butyrylcholinesterase to a far greater extent than it inhibits the abnormal enzyme. Under standardized test conditions, dibucaine inhibits expression of the normal enzyme by about 80%; the abnormal enzyme, by about 20%.

Mivacurium is also hydrolyzed by plasma cholinesterase, at a rate 70–90% of succinylcholine hydrolysis. Patients with cholinesterase variants respond similarly to mivacurium and succinylcholine. Recovery of neuromuscular function can be prolonged up to 60-fold in atypical cholinesterase homozygotes [41].

Although the dibucaine number indicates an individual's genetic makeup with respect to butyrylcholinesterase, it does not measure the concentration of the enzyme in plasma nor does it indicate the efficiency of the enzyme in hydrolyzing a substrate such as succinylcholine or mivacurium. Both of the latter factors are determined by measuring butyrylcholinesterase activity, which may be influenced by genotype. DNA analysis, although not carried out routinely, can be used to identify heterozygous carriers of atypical alleles [41].

35.3.3 Malignant Hyperthermia

Malignant hyperthermia (MH) was one of the earliest potentially fatal pharmacogenetic disorders and was first described in the 1960s. The major intracellular store of calcium (Ca^{2+}) is the sarcoplasmic reticulum (SR). The T-tubule membranes are invaginations of sarcolemma and represent approximately 80% of the plasma membrane of a skeletal muscle fiber. On excitation, the action potential generated at the neuromuscular junction spreads over the muscle fiber membrane and propagates into the interior through the T tubules. T-tubule depolarization permits the release of Ca^{2+} from the SR, which is in close association with the T tubules. T tubule–mediated SR calcium release results in what is known as excitation–contraction (EC) coupling of skeletal muscle. Thus, T-tubule membranes represent an essential element in the EC-coupling process.

EC coupling is the process by which muscle fiber depolarization initiates contraction. The action potential is transmitted to all fibrils in the fiber via the T system. It triggers the release of Ca^{2+} from the terminal cisterns—the lateral sacs of the SR next to the T system. Depolarization of the T-tubule membrane activates the sarcoplasmic reticulum via dihydropyridine receptors (DHPRs), named for the drug dihydropyridine, which blocks them. DHPRs are voltage-gated Ca^{2+} channels in the T-tubule membrane. In cardiac muscle, influx of Ca^{2+} via these channels triggers the release of Ca^{2+} stored in the SR (calcium-induced calcium release) by activating the ryanodine receptor (RyR).

The RyR is named after the plant alkaloid ryanodine, which was used in its discovery. It is a ligand-gated Ca^{2+} channel with Ca^{2+} as its natural ligand. In skeletal muscle, Ca^{2+} entry from the extracellular fluid (ECF) by this route is not required for Ca^{2+} release. Instead, the DHPR that serves as the voltage sensor allows the release of Ca^{2+} from the nearby SR via physical interaction with the RyR. The released Ca^{2+} is quickly amplified through calcium-induced calcium release. Ca^{2+} is reduced in the muscle cell by the sarcoplasmic or endoplasmic reticulum Ca^{2+} ATPase (SERCA) pump.

The SERCA pump uses energy from ATP hydrolysis to remove Ca^{2+} from the cytosol back into the terminal cisterns, where it is stored until released by the next action potential. Once the Ca^{2+} concentration outside the reticulum has been lowered sufficiently, chemical interaction between myosin and actin ceases and the muscle relaxes. Note that ATP provides the energy for both contraction (at the myosin head) and relaxation (via SERCA). If Ca^{2+} transport into the reticulum is inhibited, relaxation does not occur, even though there are no more action potentials; the resulting sustained contraction is called a contracture [64]. (See Figure 35.2.)

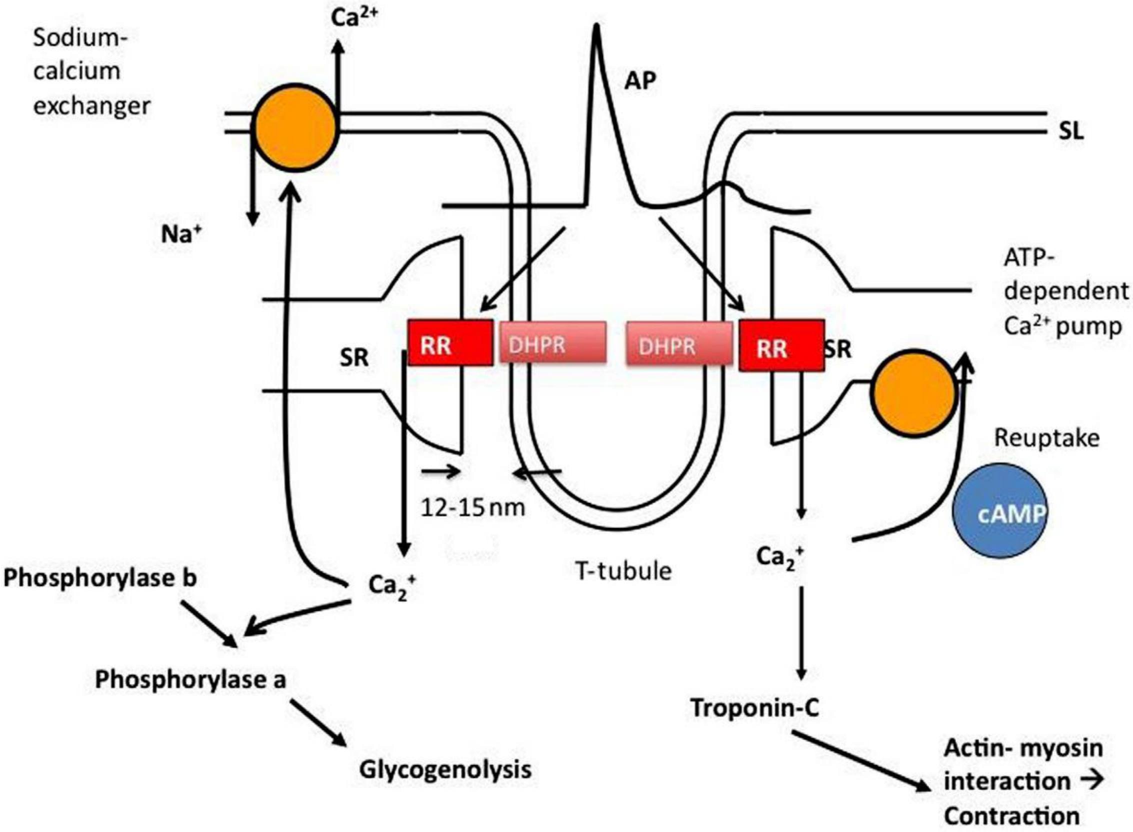


FIGURE 35.2 Calcium is released directly from the sarcoplasmic reticulum (SR) by the action potential (AP), through the ryanodine receptor (RR). Increased intracellular Ca^{2+} removes inhibition of the contractile elements by troponin, and the muscle contracts. The sodium–calcium exchanger on the sarcolemma (SL) and the cyclic AMP AMP– (cAMP-) regulated pump on the SR remove calcium from the myofibril, terminating contraction. Ca^{2+} simultaneously activates glycogenolysis to provide energy for contraction. The T tubules contain the highest concentration of dihydropyridine binding sites in the body. The dihydropyridine receptors (DHPRs) of skeletal muscle do not function as Ca^{2+} channels, but instead act as voltage sensors, transducing the sarcolemmal action potential into the release of Ca^{2+} from the SR.

MH has been described in patients of all racial groups. The majority of MH reactions occur in children and young adults, with a 2:1 male preponderance. In the United Kingdom, the MH Investigation Unit estimated that an MH reaction occurs with 1 in 20,000 to 1 in 80,000 anesthetics, depending on patient age. However, this may be a conservative estimate, as Ording reported that MH reactions occur in 1 in 15,000 anesthetics in Denmark [65]. The European MH group consensus estimate for the prevalence of genetic susceptibility to MH is about 1 in 10,000 [66].

The recognized modes of inheritance for MH susceptibility are autosomal dominant with reduced penetrance and variable expression, autosomal recessive, or multifactorial and unclassified. Reduced penetrance means that fewer offspring are affected than would be predicted by a perfectly dominant pattern. Variable expression refers to a difference of susceptibility between families with little variation within the same family. The gene for MH is located on human chromosome 19, which is also the genetic coding site for the calcium release channels of skeletal muscle SR (ryanodine receptor RyR1 locus). It is presumed that a defect in the calcium release channels results in MH susceptibility. The presence of RyR1 mutation in patients does not always relate to a positive *in vitro* contracture in response to halothane. MH in humans is not a single disease but a syndrome with multiple sites of causation and multiple mutations at them [64,67].

Even the presentation of MH is not the same in all patients. Some 30% of MH patients have had up to three uneventful anesthetics. A spectrum can occur from minor reactions to rapid temperature increase, muscle rigidity, acidosis, arrhythmias, and death. Some reactions have greater latency to onset and may not manifest until the postoperative period. MH may not always occur in response to triggering agents. Thus, there may be different genes causing MH in different families, or there may be other predisposing factors that are expressed differently in patients or families. MH may be described as a heterogeneous genetic disorder with a highly variable clinical presentation. There are no MH-specific clinical features, and the diagnosis depends on knowledge of features that can occur during an episode. MH is characterized by signs and symptoms of hypermetabolism (up to 10 times normal).

The earliest signs of MH are related to enormous increases in the patient's metabolic rate, reflecting the ability of triggering drugs to cause an imbalance in calcium homeostasis in skeletal muscle cells. In some patients, metabolic stimulation is evident clinically within 10 minutes of administration of volatile anesthetics, whereas in others several hours may elapse. Hypermetabolism leads to increased carbon dioxide production with associated tachypnea. In addition, lactic acidosis develops, and the presence of mixed respiratory and metabolic acidosis stimulates sympathetic nervous system activity with associated tachycardia. Increased carbon dioxide production occurs early, emphasizing the value of continuous capnography. Temperature rise may be a late sign, but core temperature may increase as early as 15 minutes after exposure to a triggering agent. Cardiac dysrhythmias, such as ventricular bigeminy, multifocal ventricular premature beats, and ventricular tachycardia, may also occur, especially with hyperkalemia resulting from rhabdomyolysis and sympathetic nervous system stimulation. Other complications are hemolysis, myoglobinuria, rhabdomyolysis, and renal failure [64,66,68].

Successful treatment of MH depends on early diagnosis and a preplanned therapeutic regimen. Treatment can be categorized as causal or symptomatic. Causal treatment is directed at correcting the underlying causative mechanisms. Symptomatic treatment is directed toward maintaining renal function and correcting hyperthermia, acidosis, and arterial hypoxemia.

Once MH is recognized it is imperative to act quickly: (1) discontinue the inhalational agent, (2) hyperventilate with 100% oxygen and administer dantrolene, (3) cool the patient and treat symptoms. Dantrolene, administered intravenously, is the only drug that is reliably effective in treating MH. It decreases the release of calcium from the SR, restoring the balance between release and reuptake. Active cooling may entail surface cooling and intracavitary lavage of the stomach and bladder using cold saline solutions infused through peripheral intravenous catheters; they should also be cooled. Cooling is discontinued when body temperature decreases to 38°C.

Other symptomatic therapy includes intravenous administration of sodium bicarbonate to correct metabolic acidosis and hyperkalemia, hydration with saline, and maintenance of urine output using osmotic or tubular diuretics. Intravenous administration of glucose-insulin helps drive potassium intracellularly and provides an exogenous energy source with which to replace depleted cerebral metabolic substrates [64,66,68].

More than 100 variants of RyR1 have been reported, although the RyR1 gene mutation on chromosome 19 is predominant. Of variants that are predicted to lead to an amino acid chain in the ryanodine receptor protein, 50 have been detected in UK MH families. More than 20 mutations have been functionally characterized in experimental systems—an essential step in identifying a mutation as potentially MH causative. These functional mutations have the potential to be used in diagnosis. However, in a minority of families (approximately 10–20%), no RyR1 defect can be found, even after coding the entire sequence of the gene [64,69].

Genetic linkage studies have implicated five other chromosomal candidate loci, but a mutation has been found in only one of them (the alpha-1 subunit of the dihydropyridine receptor), and in this case only two mutations in two different families have been reported worldwide [70,71]. The likelihood that defects in other genes contribute to MH susceptibility is high [72]. The inability to find an RyR1 mutation does not exclude the possibility of MH even when a mutation found in other members of the family is not present. An individual who does not have the familial mutation, however, has an approximately 5% chance of actually being susceptible to MH; his or her status must therefore be confirmed by *in vitro* contracture testing before being administered MH-triggering drugs [66–68,73,74].

Determining MH linkage with other diseases has been problematic—only central core disease appears to be truly linked, and most families with central core disease have a mutation at the RyR1 locus [74]. Other syndromes associated with multicore myopathy are Evan's myopathy and the King-Denborough syndrome/phenotype. In Duchenne's muscular dystrophy, the balance of opinion has shifted from an association with MH to an one with anesthesia-induced rhabdomyolysis.

The contracture halothane caffeine test is the most common method of diagnosing MH. This test requires muscle biopsy and measures contracture in response to caffeine and halothane. The test is very sensitive, meaning that most patients with a risk of and susceptibility to MH will be identified. Problems with it include the requirement for a fresh sample at the testing site, the invasiveness and expense of the test, and limited biopsy centers. MH centers offer patients (and their families) genetic testing if they have a positive result from the caffeine halothane contracture test; the centers screen for the 17 most common RyR1 mutations. Although the test is very specific, only about 25% of individuals at risk for MH are detected because of the multiple areas of MH mutation [66,68,75].

Mortality from MH has decreased from 80% thirty years ago to less than 5% today. In the future, DNA testing will replace the *in vitro* contracture test for high-risk individuals. When DNA testing achieves 100% sensitivity and specificity, it will be suitable for population screening, but this requires that all gene defects contributing to MH susceptibility be identified [4,64,76,77].

35.3.4 Pharmacogenetics of Specific Drugs in Anesthesia

The following subsections focus on drugs directly related to the administration of anesthesia related to anxiolysis, analgesia, and muscle relaxation, which are integral to anesthesia administration (See Tables 35.1–35.3). Other drugs that the anesthetist may come across are discussed elsewhere in this book.

35.3.4.1 Benzodiazepines

Benzodiazepeines (BDZs) are commonly used preoperatively as anxiolytic agents for control of seizures. Their administration is concomitant with induction of anesthesia agents. Most benzodiazepines are metabolized by hepatic cytochrome P450 enzymes to more polar metabolites and are excreted in bile and urine. The relative antianxiety potencies of the different BDZs correlate with their relative potencies as agonists at the BDZ receptor. The resultant inhibitory effects in the brain are responsible for BDZs' anti-anxiety, sedative, and anticonvulsant effects. The inhibition of the spinal pathways results in skeletal muscle relaxation. The therapeutic use of particular BDZs is as much related to their potency and solubility as to their inherent pharmacokinetic differences. Given orally at normal therapeutic doses, they have little effect on cardiovascular, respiratory, or autonomic function. Respiratory depression and reduced systolic blood pressure may occur, but these are seen principally with intravenous administration or overdose.

Temazepam, midazolam, and diazepam are the preferred drugs for preoperative anxiolysis because they are rapidly absorbed by the oral route and short-acting after single doses. The choice of drug varies according to local practice and to the availability of oral preparations. Lorazepam is also used for preoperative anxiolysis, particularly when it is desirable to dampen sympathetic outputs before major cardiac surgery. In these circumstances, lorazepam may be given the evening before the operation as well as in the morning.

TABLE 35.1 Major CYP Enzymes and their Substrates in Anesthetic Practice

CYP2B6	CYP2C9	CYP2C19	CYP2D6	CYP2E1	CYP3A4	CYP3A5
Propofol	Propofol	Diazepam	Codeine	Sevoflurane	Diazepam	Diazepam
	Parecoxib	Barbiturates		Halothane	Midazolam	
		Omeparzole		Isoflurane	Temazepam	
		Pantoprazole		Paracetamol	Fenatanyl	
					Alfentanil	
					Lignocaine	
					Vecuronium	

Source: Data from Bhasker et al. [48] and from Indiana University [17].

TABLE 35.2 Abnormalities of Butyrylcholinesterases

Patient	Genotype	Incidence	Dibucaine number	Block Duration with Mivacurium and Suxamethonium
Homozygous typical	E1u: E1u	96%	>70	No change
Heterozygous atypical	E1u: E1a	1:500	40–70	20–30 minutes
Homozygous atypical	E1a: E1a	1:3200	<20	4–8 hours
Homozygous fluoride-resistant	E1f: E1f	1:32,000	70–80	1 hour
Homozygous silent	E1s: E1s	1:100,000	-	>6 hours

Source: Data from Jensen et al. [37], Staron [39], and Kalow [40].

TABLE 35.3 Example Genotypes Linked to Individual Differences in Metabolism, Efficacy, and Side Effects of Commonly Used Anesthetic Agents

Drug	Gene	Effects
Volatile anesthetics		
Halothane, other inhalational anesthetics	CYP2E1	Hepatitis
Halothane, isoflurane	RyR1	Malignant hyperthermia syndrome
Muscle relaxants		
Succinylcholine, mivacurium	Plasma butyrylcholinesterase	Prolonged muscle relaxation
Succinylcholine	RyR1	Malignant hyperthermia syndrome
Benzodiazepines		
Diazepam	CYP2C19	Increased half-life, prolonged sedation, greater likelihood of unconsciousness
Midazolam	CYP3A4	Reduced clearance of systemic (but not oral) midazolam
Midazolam	CYP3A5	Reduced clearance of systemic (but not oral) Midazolam
Opioids		
Morphine, morphine-6-glucuronide	μ opioid receptor	Reduced nausea, vomiting, sedation, drowsiness
Morphine	Uridine diphosphate glycosyl transferase	More rapid glucuronidation
Codeine	CYP2D6	Slower conversion to morphine; reduced analgesic effect (but no adverse effects)
Tramadol	CYP2D6	Slower conversion to O- and N-demethyltramadol; greater need for rescue analgesics after surgery
Methadone	CYP2D6	Slower metabolic clearance (although not consistently)
NSAIDs and Others		
Celecoxib, naproxen, piroxicam, ibuprofen, flurbiprofen	CYP2c9	Slower metabolism
Paracetamol	CYP2e1	Accelerated elimination rate
Acetaminophen	Tumor necrosis factor β	Lower risk of encephalopathy in patients with acetaminophen-induced acute liver failure
Various NSAIDs	HLA-DRB1	Higher risk of anaphylactoid reaction

Source: Adapted from Palmer et al. [3].

CYP—cytochrome P450; HLA—human leukocyte antigen; NSAID—nonsteroidal anti-inflammatory drug; RyR1—ryanodine receptor.

The BDZ of choice for intravenous sedation is midazolam, the only water-soluble benzodiazepine. It is preferred for sedation in intensive care patients and is particularly useful when given in combination with other drugs, such as alfentanil and propofol. Although the duration of action after bolus injection is short due to rapid redistribution, when administered to critically ill patients as a continuous infusion for longer than 24 hours, its clinical effect is often prolonged due to accumulation and possibly also due to α-hydroxymidazolam, midazolam's active metabolite.

Alphahydroxymidazolam is 60–80% as potent as midazolam, with an elimination half-life of about 1 hour in patients with normal renal function. The clinical effects of midazolam may be increased in patients with renal failure owing to α-hydroxymidazolam accumulation. Diazepam is less suitable as an infusion on account of its longer half-life and the even longer half-life of its principal metabolite. Given intravenously, both diazepam and midazolam are effective first-line treatments for status epilepticus. For intractable status epilepticus, clonazepam, a longer-acting alternative, which can also be given by intravenous infusion, is often used.

35.3.4.2 Diazepam

The half-life of diazepam in individuals homozygous for the A allele of the cytochrome CYP2C19 G681A polymorphism is four times longer than individuals homozygous for the major (G) allele, presumably because of markedly decreased CYP2C19 metabolic activity. Individuals heterozygous for the A allele have a diazepam half-life somewhere between these two extremes. These genetic differences may manifest clinically as prolonged sedation or unconsciousness after diazepam administration [78,79].

35.3.4.3 Midazolam

Midazolam is metabolized by two different enzymes (CYP 3A4, CYP3A5) present in both the liver and the intestine. So far, there is no evidence that polymorphisms in either of these enzymes cause any effect on drug action when given orally. However, coadministration of midazolam and fentanyl can prolong the action of fentanyl due to inhibition of fentanyl metabolism by midazolam's competing metabolism at the liver enzyme CYP3A4 [80].

35.3.5 Inhalational Anesthetic Agents

Volatile anesthetics or inhalational agents are used mostly for maintenance and occasionally for induction of anesthesia. The commonly used inhalational anesthetic agents in current practice are isoflurane, sevoflurane, and desflurane. Halothane is still used in developing countries, but in the developed world it has been superseded by the newer volatile anesthetics just mentioned. Two main pharmacogenetic conditions of particular relevance to anesthesia are malignant hyperthermia and halothane- (and sevoflurane-) induced hepatitis.

Individual sensitivity to volatile anesthetics based on genetics has not been well studied or defined. Based on research by Liem et al., we know that many redheads have distinct polymorphisms in the melanocortin-1 receptor that make them less sensitive to desflurane (requiring more drug to achieve the mean alveolar concentration, or MAC) [81]. Other than this study, there is little to show that any genetic variability causes a change in sensitivity to individual volatile agents.

Halothane-induced hepatitis, which has been widely studied, seems to result from an immune response to halothane metabolites processed by the cytochrome enzyme CYP2E1. Halothane-induced hepatitis occurs in approximately 1 in 10,000 patients. Although its tendency to occur within families suggests that it is hereditary, the genetic mechanisms involved are not entirely clear. This is because CYP2E1 activity depends as much on body mass, diet, alcohol consumption, and age as it does on any polymorphisms discovered to date. Mechanisms of halothane toxicity are idiosyncratic reaction, immuno-allergic type; the protein adducts formed in the initial toxic reaction provide the hapten for the formation of antibodies, which augments damage on re-exposure. Covalent binding of halothane metabolites to tissue proteins has been well described. Up to one-third of inhaled halothane may be oxidatively metabolized by CYP2E1 and 2A6.

Trifluoroacetyl chloride, which has a strong ability to acetylate liver proteins, is formed during the synthesis of trifluoroacetic acid (TFA). Only a small amount of halothane (~%1) is metabolized through the reductive pathway by CYP2A6 and 3A4. This pathway is favored under hypoxic conditions, resulting in the release of bromide and fluoride ions and in the formation of other volatile organic metabolites. Several drugs or chemicals have been shown to induce hepatitis with autoimmune involvement. Adduct formation of an activated metabolite is believed to act as a trigger and to induce a specific immune response.

Two types of halothane-mediated hepatotoxicity have been defined. The first, type I, occurs postoperatively and is mild and self-limiting, with a mild form of hepatocellular injury that can be observed in about 20% of patients. The mild hepatic injury is assumed to result from the direct action of halothane on liver cells. Two clinically detectable factors appear to contribute to this. The first is a transient elevation of liver enzymes and the second is alteration of cellular integrity, which can be detected by electron microscopy. Lesions result from intracellular degradation of halothane via its anaerobic and aerobic pathways, in combination with local hypoxia caused by an alteration of the hepatic oxygen demand–supply relationship.

The second type of halothane-mediated hepatotoxicity is type II—halothane hepatitis. The incidence of type II hepatotoxicity after halothane administration is one in 10,000–30,000 adult patients. Its probable mechanism is an immune-mediated hepatotoxicity, in which antibodies are directed against modified liver microsomal proteins on hepatocyte surfaces. There is strong evidence that the fulminant form of halothane hepatitis is mediated by the patient's own immune system. Besides signs of mild cellular injury, tissue acetylation is usually found due to the generation of reactive intermediates from halothane metabolism. The acetylation of intacellular proteins is considered the first step in the pathogenesis of the severe type of hepatic injury. The second step involves the formation of antibodies directed toward these acetylated neo-antigens.

The consequence of the immune response with regard to hepatic injury is severe and sometimes fatal. CYP2E1 is a major catalyst in the formation of trifluoroacetylated proteins, which have been implicated as target antigens in the mechanism of halothane hepatitis. Patients with halothane hepatitis or severe hepatic damage caused by other halogenated anesthetics produce metabolites against several liver trifluoroacetylated microsomal proteins [82].

There are reports of hepatoxicity following exposure to sevoflurane. Sevoflurane is a widely used volatile anesthetic agent in both adult and pediatric patients, with rapid onset and washout. Because of its rapid onset of action and lack of irritability to the airways, sevoflurane can be used for both induction and maintenance of anesthesia. Prospective, serial blood tests have often demonstrated minor transient elevations in serum aminotransferase (ALT) levels 1–2 weeks following major surgery.

Clinically apparent, severe hepatic injury from sevoflurane is very rare, with only isolated case reports having been published. The injury is marked by acute elevations in serum ALT levels (5–50-fold) and the appearance of jaundice 2–21 days after surgery. There are usually minimal increases in alkaline phosphatase and gammaglutamyl transpeptidase levels. Jaundice is usually preceded by a day or two of fever and may be accompanied by rash and eosinophilia.

Acute liver injury may be self-limiting and resolve in 4–8 weeks, but it can also be severe and associated with acute liver failure. A strong risk factor is previous exposure to any of the halogenated anesthetics and particularly a history of halothane hepatitis or unexplained fever and rash after anesthesia with one of these agents. The differential diagnosis of acute liver injury after surgery and anesthesia is sometimes difficult, and a clinical picture similar to sevoflurane-induced hepatitis can be caused by shock or ischemia, sepsis, other idiosyncratic forms of drug-induced liver injury, and acute viral or herpes hepatitis [83–85].

The mechanism of sevoflurane hepatotoxicity is suspected to be similar to that of halothane and to be associated with creation of reactive intermediates. Sevoflurane is metabolized, to a small extent, by the microsomal drug-metabolizing enzyme CYP 2E1 to a TFA-reactive intermediate that is capable of binding to multiple intracytoplasmic proteins, forming potentially immunogenic adducts. The TFA adducts induce antibodies that can be detected in patients with sevoflurane as well as halothane hepatotoxicity and are also found in a proportion of healthcare workers exposed to the volatile anesthetics [84,85].

Among approximately 50,000 liver transplants performed in the United States between 1990 and 2002, 137 [0.2%] were because of idiosyncratic drug-induced acute liver failure. Three of these were attributed to halothane, one to isoflurane, and none to other halogenated anesthetics.

35.3.6 Neuromuscular Blocking Agents

Most of the commonly used neuromuscular blocking drugs (NMBDs) are quaternary ammonium compounds. They are structurally related to acetylcholine (ACh), which contains a quaternary nitrogen group ($N^+(CH3)_3$). As with ACh, the positive nitrogen atoms of NMBDs are attracted to the α-subunits of the postsynaptic nicotinic receptor.

35.3.6.1 Depolarizing Muscle Relaxants

Depolarizing muscle relaxants (DMRs) are agonists located at ACh receptors. Succinylcholine is the only DMR in clinical use. Recovery from neuromuscular block occurs as succinylcholine diffuses away from the neuromuscular junction and down a concentration gradient in association with decline in its plasma levels. In the plasma, it is metabolized by butyrylcholinesterase. As discussed earlier, reduced/absent butyrylcholinesterase activity, a result of inherited or acquired factors, may alter the duration of action of succinylcholine and lead to prolonged paralysis [1,40,41].

35.3.6.2 Nondepolarizing Muscle Relaxants

Nondepolarizing muscle relaxants (NMDRs) antagonize the action of ACh in a competitive manner at the postsynaptic nicotinic receptor. Unlike depolarizing drugs, they do not produce conformational changes in the receptor. Binding to one or both α-subunits prevents access by ACh to depolarize the receptor. The binding of antagonists to receptors is dynamic, with repeated association and dissociation. If the concentration of ACh is increased, it has a higher chance than the antagonist of occupying the receptor sites. Except for mivacurium, which is metabolized by butyrylcholinesterase, no genetic factors have been implicated in prolonged duration of NMDR action. Prolonged duration of action results from reduced or absent butyrylcholinesterase activity [41,86,87].

35.3.7 Opioids

Opioids—the most common drugs used by anesthetists worldwide for analgesia and sedation in perioperative and intensive-care settings—exert their pharmacological effects by binding to specific cellular receptors. At least three distinct receptor subtypes (μ, κ, and δ) have been identified and studied extensively; they are primarily responsible for opiods' observed pharmacological effects. Binding to (or blockade of) these receptors alters many clinically important physiological functions. Most of the clinical effects observed are mediated by the μ receptor.

A newer receptor known as nociceptin/orphanin FQ peptide (NOP) and variants of μ-receptor genes have been identified; however, their impact on pharmacotherapy requires further research and understanding. Opioid action is very difficult to characterize by single polymorphisms.

The pharmacogenomics of opioid action can be explained by categorizing the variables into receptor polymorphism, drug transport polymorphism, and metabolism [4].

35.3.7.1 μ Receptor Polymorphism

The μ receptor is the principal site of opioid action. Its two common variants, present in 5–10% of Caucasians, are A118G and the promotor polymorphism G-172T. Although the G118 allele has no apparent relation to a patient's analgesic response to morphine (nor does its active metabolite morphine-6-glucouronide), preliminary data suggest that this allele is associated with reduced severity of opiod adverse effects, which include nausea, vomiting, pupil dilation, and sedation. Therefore, G118 allele carriers may tolerate higher opiate doses than noncarriers [51–53,55,88–91].

35.3.7.2 Drug Transport Polymorphisms

Morphine, methadone, fentanyl, and meperidine are all transported (e.g., across cell membranes, to the brain) by the ABC transporters, specifically the ABCB1 efflux transporter. Research on this transporter is just beginning, although investigations carried out so far show that patients with two copies of the relevant polymorphisms may require as much as a 2-fold increase in opioid dosage [8,12,13,91].

35.3.7.3 Metabolism

Opioid metabolism by cytochrome P450 enzymes and by enzymes that regulate glucuronidation to active metabolites influence the concentration of drug available at the effect site and therefore its clinical efficacy [91]. Genes encoding enzymes responsible for these metabolic pathways exhibit considerable polymorphism, which results in wide variability in interindividual responses to some opioids. The clinical effects of the weaker opioids, codeine, dihydrocodeine, tramadol, oxycodone, and hydrocodone, rely on the formation of more potent hydroxyl metabolites (such as morphine, dihydromorphine, and oxymorphone) via a metabolic pathway mediated by the cytochrome P450 enzyme CYP2D6.

The gene encoding CYP2D6 exhibits significant polymorphism, with 100 variant alleles identified. Variant allele expression results in a number of clinical phenotypes: poor metabolizers express two nonfunctional alleles (e.g., 4*, 5*, and 6*), intermediate metabolizers express at least one reduced functional allele (e.g., 9*, 10*, and 41*), and extensive metabolizers express at least one functional allele [22]. A fourth phenotype, the ultrarapid metabolizer, develops when a gene duplication event results in the individual having multiple copies of the functional allele. Approximate frequencies of these phenotypes for the Caucasian population are poor metabolizers 5–10%, intermediate metabolizers 10–15%, extensive metabolizers 65–80%, and ultrarapid metabolizers 5–10%. The prevalence of variant alleles differs considerably according to ethnicity—for example, the frequency of CYP2D6 gene duplication associated with ultrarapid metabolism ranges from 0.5% in China to 29% in Ethiopia [16,22,91,92].

35.3.7.4 Morphine

Multiple liver cytochromes influence the morphine metabolism. Morphine glucoronidation is the major path of first-pass metabolism. An alteration in the genotype of UDGT causes increased glucoronidation. Surgical patients homozygous for the C161T and C802T UDGT gene polymorphisms (which are always inherited together) demonstrate more rapid morphine glucuronidation than do patients who are heterozygous or wild type, therefore causing a decrease in morphine effectiveness. Most recently, it was found that CYP2D6 poor and ultrarapid metabolizers have very poor analgesia with morphine when compared to extensive and intermediate metabolizers [2,51,58,93,94].

35.3.7.5 Codeine

Codeine has a minimal clinical effect until it is metabolized to morphine in the liver by CYP2D6 (about 10% of the drug) [95–97]. Approximately 5–10% of Caucasian patients (poor metabolizers) are unlikely to gain full benefit from codeine administration, although they are just as likely to suffer codeine-related side effects [98]. In children, this number may be as high as 36% [99]. The effect may be compounded by low drug absorption.

There is a small population of individuals, most commonly those of East African descent, who are ultrarapid metabolizers and convert the majority of codeine into morphine. For these patients, respiratory depression and apnea are very real possibilities after even a single dose. This is of particular concern in nursing mothers [100]. Interestingly, the severity of codeine's adverse effects does not seem to be associated with the CYP2D6 genotype, thus explaining how codeine can impair driving ability even when the driver's blood does not contain detectable morphine concentrations [25].

35.3.7.6 Methadone

CYP2D6 poor metabolizers have decreased methadone clearance, although this decrease is not always consistent [101]. It may be secondary to the influence of CY2B6 and CY3A4 in methadone metabolism. CYP3A4 is important in the metabolism of several other opiates, such as fentanyl, alfentanil, and sufentanil. Interestingly, CYP3A4-mediated metabolism is 40% more rapid in females than in males. The clinical implication is that females have a 3-fold higher incidence of awareness during total intravenous anesthesia and have a more rapid awakening [50,102].

35.3.7.7 Tramadol

Tramadol is an opiate that is metabolized to O- and N-demethyltramadol by the CYP2D6 enzyme system. Poor metabolizers maintain high tramadol levels for an extended period of time. Because this actually results in a diminishment of drug effectiveness, the percentage of tramadol nonresponders is significantly higher in the poor metabolizer group (46.7%) compared with the normal (extensive metabolizer) group (21.6%). In contrast, CYP2D6*1 allele carriers were found to be twice as likely to require additional postoperative analgesics because of increased tramadol metabolism [50,103–105].

35.3.8 Nonsteroidal Anti-Inflammatory Drugs

Nonsteroidal anti-inflammatory drugs (NSAIDs) are commonly used to treat various types of pain. Their mode of action is inhibition of the cyclo-oxygenase enzyme system. Many NSAIDs are metabolized by the cytochrome enzyme CYP2C9. Celecoxib, naproxen, piroxicam, ibuprofen, and flurbiprofen are more slowly metabolized in individuals with CYP2C9*3 than in those without it [98,106,107]. This reduction in metabolism tends to be greater in *3 homozygotes than in heterozygotes, which suggests that *3 carriers may experience either more clinical benefit or greater adverse effects from NSAIDs, or both. However, no studies have yet confirmed this. In contrast, metabolism of diclofenac seems largely unrelated to the CYP2C9 genotype [28,29].

Preliminary evidence indicates that the *11 allele of the gene for the human leukocyte antigen HLA-DRB1, which is involved in immune response, may be associated with anaphylactoid reactions to various NSAIDs. However, there is no evidence to suggest any association between NSAID activity and cyclooxygenase-2 gene polymorphisms. This is surprising, given that cyclooxygenase-2 inhibition is the NSAID mechanism of action. It has been suggested that the cyclooxygenase-2 enzyme is so important for maintaining body homeostasis that any polymorphisms that might significantly alter its function have been suppressed by natural selection. This seems an unlikely explanation because the G-765C polymorphism of the cyclooxygenase-2 gene was recently shown to be independently associated with a decreased risk of myocardial infarction and stroke.

35.3.8.1 Paracetamol

In therapeutic doses, paracetamol (acetaminophen) is one of the most effective and well-tolerated drugs in adults and children. It has been in use in clinical practice for more than 50 years and has been the cornerstone of pain and fever management. Although paracetamol has been used for the last century, its mechanism of action is still unknown; however, it is generally considered to be a weak inhibitor of prostaglandin (PG) synthesis. *In vivo*, paracetamol's effects are similar to those of selective COX-2 inhibitors. Although it decreases PGs, it does not reduce inflammation in conditions like rheumatoid arthritis, as the COX-2 inhibitors do.

Paracetamol metabolism involves CYP2E1. Individuals who are homozygous for the CYP2E1 c2 allele (which lacks a DNA segment that limits protein manufacture in the wild-type c1 allele) metabolize paracetamol twice as fast as do heterozygous or wild-type individuals. Some data imply that a mutant allele (B2) of the tumor necrosis factor β gene protects against encephalopathy in patients with paracetamol-induced acute liver failure [18,108].

35.3.8.2 Antiemetics

Ondansetron is a serotonin (5HT3) receptor antagonist used for the prevention and treatment of postoperative nausea and vomiting. Ultrarapid metabolizers of CYP2D6 have a high incidence of ondansetron failure, caused by their rapid metabolism of ondansetron and low overall drug exposure [19,109].

35.4 CONCLUSION

Even though anesthesia was at the forefront of the pharmacogenetic discoveries, applied pharmacogenetics in anesthesia is still in its infancy. Nevertheless, the range of anesthetic drugs along with the fact that the majority of subjects undergoing elective anesthesia do not have comorbidities and that those who do are generally at high risk and are exposed to multiple pharmacological agents—all of these factors indicate that pharmacogenomics and personalized medicine are the way forward in tailoring interventions to individual needs as a way to minimize adverse events [110].

REFERENCES

[1] Kalow W, Gunn DR. The relation between dose of succinylcholine and duration of apnea in man. J Pharmacol Exp Ther 1957;120:203–14.

[2] Landau R, Bollag LA, Kraft JC. Pharmacogenetics and anaesthesia: the value of genetic profiling. Anaesthesia 2012;67:165–79.

[3] Palmer SN, Giesecke NM, Body SC, Shernan SK, Fox AA, Collard CD. Pharmacogenetics of anesthetic and analgesic agents. Anesthesiology 2005;102:663–71.

[4] Searle R, Hopkin PM. Pharmacogenomic variability and anaesthesia. Br J Anaesth 2009;103:14–25.

[5] FDA, Drug Safety Communication: Safety Review Update of Codeine Use in Children; New Boxed Warning and Contraindication on Use After Tonsillectomy and/or Adenoidectomy. 2013 <www.fda.gov/downloads/Drugs/DrugSafety/UCM339116.pdf>.

[6] Ambudkar SV, Kimchi-Sarfaty C, Sauna ZE, Gottesman MM. P-glycoprotein: from genomics to mechanism. Oncogene 2003;22:7468–85.

[7] Marzolini C, Paus E, Buclin T, Kim RB. Polymorphisms in human MDR1 (P-glycoprotein): recent advances and clinical relevance. Clin Pharmacol Ther 2004;75:13.

[8] Ieiri I, Takane H, Otsubo K. The MDR1 (ABCB1) gene polymorphism and its clinical implications. Clin Pharmacokinet 2004;43:553–76.

[9] Leschziner GD, Andrew T, Pirmohamed M, Johnson MR. ABCB1 genotype and PGP expression, function and therapeutic drug response: a critical review and recommendations for future research. Pharmacogenomics J 2007;7:154–79.

[10] Pauli-Magnus C, Kroetz DL. Functional implications of genetic polymorphisms in the multidrug resistance gene MDR1 (ABCB1). Pharm Res 2004;21:904–13.

[11] International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61.

[12] Wandel C, Kim R, Wood M, Wood A. Interaction of morphine, fentanyl, sufentanil, alfentanil and loperamide with the efflux drug transporter P-glycoprotein. Anesthesiology 2002;96:913–20.

[13] Sweeney BP. Pharmacogenomics and anaesthesia: explaining the variability in response to opiates. Eur J Anaesthesiol 2007;24:209–12.

[14] Nebert DW, Russel DW. Clinical importance of the cytochrome P450. Lancet 2002;360:1152–62.

[15] Wilkinson GR. Drug metabolism and variability among patients in drug response. N Engl J Med 2005;352(21):2211–21.

[16] Ingelman-Sundberg M. Pharmacogenomics of cytochrome P450 and its applications in drug therapy: the past, present and future. Trends Pharmacol Sci 2004;25(4):193–200.

[17] Indiana University, P450 Drug Interaction Table, School of medicine website. 2013 <http://medicine.iupui.edu/clinpharm/ddis/main-table/>.

[18] Ueshima Y, Tsutsumi M, Takase S, Matsuda Y, Kawahara H. Acetaminophen metabolism in patients with different cytochrome P-4502E1 genotypes. Alcohol Clin Exp Res 1996;20:25A–8A.

[19] Candiotti KA, Birnbach DJ, Lubarsky DA, Nhuch F, Kamat A, Koch WH, et al. The impact of pharmacogenomics on postoperative nausea and vomiting: do CYP2D6 allele copy number and polymorphisms affect the success or failure of ondansetron prophylaxis? Anesthesiology 2005;102:543–9.

[20] Eichelbaum M, Kroemer HK, Fromm MF. Impact of P450 genetic polymorphism on the first-pass extraction of cardiovascular and neuroactive drugs. Adv Drug Deliv Rev 1997;27:171–99.

[21] Kirchheiner J, Schmidt H, Tzvetkov M, Keulen JT, Lötsch J, Roots I, et al. Pharmacokinetics of codeine and its metabolite morphine in ultra-rapid metabolizers due to CYP2D6 duplication. Pharmacogenomics J 2007;7:257–65.

[22] Zanger UM, Raimundo S, Eichelbaum M. Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. Naunyn-Schmiedebergs Arch Pharmacol 2004;369:23–37.

[23] Bertilsson L. Geographical/interracial differences in polymorphic drug oxidation: current state of knowledge of cytochromes P450 (CYP) 2D6 and 2C19. Clin Pharmacokinet 1995;29:192–209.

[24] Ma JD. Genetic polymorphisms of cytochrome P450 enzymes and the effect on interindividual, pharmacokinetic variability in extensive metabolizers. J Clin Pharmacology 2004;44:447–56.

[25] Eckhardt K, Li S, Ammon S, Schanzle G, Mikus G, Eichelbaum M. Same incidence of adverse drug events after codeine administration irrespective of the genetically determined differences in morphine formation. Pain 1998;76:27–33.

[26] Bachs L, Skurtveit S, Morland J. Codeine and clinical impairment in samples in which morphine is not detected. Eur J Clin Pharmacol 2003;58:785–9.

[27] Wadelius M, Pirmohamed M. Pharmacogenetics of warfarin: current status and future challenges. Pharmacogenomics J 2007;7:99–111.

[28] Brenner SS, Herrlinger C, Dilger K, Murdter TE, Hofmann U, Marx C, et al. Influence of age and cytochrome P450 2C9 genotype on the steady-state disposition of diclofenac and celecoxib. Clin Pharmacokinet 2003;42:283–92.

[29] Yasar U, Eliasson E, Forslund-Bergengren C, Tybring G, Gadd M, Sjöqvist F, et al. The role of CYP2C9 genotype in the metabolism of diclofenac *in vivo* and *in vitr*o. Eur J Clin Pharmacol 2001;57:729–35.

[30] Andersson T. Omeprazole drug interaction studies. Clin Pharmacokinet 1991;21:195–212.

[31] Humphries TJ. Clinical implications of drug interactions with the cytochrome P-450 enzyme system associated with omeprazole. Dig Dis Sci 1991;36:C1665–9.

[32] Zubieta JK, Heitzeg MM, Smith YR, Bueller JA, Xu K, Xu Y, et al. COMT val(158)met genotype affects mu-opioid neurotransmitter responses to a pain stressor. Science 2003;299:1240–3.

[33] Rakvag TT, Klepstad P, Baar C, Kvam TM, Dale O, Kaasa S, et al. The Val158Met polymorphism of the human catechol-O-methyltransferase (COMT) gene may influence morphine requirements in cancer pain patients. Pain 2005;116:73–8.

[34] Kim H, Lee H, Rowan J, Brahim J, Dionne RA. Genetic polymorphisms in monoamine neurotransmitter systems show only weak association with acute post-surgical pain in humans. Mol Pain 2006;2:24.

[35] Hosokawa M. Structure and catalytic properties of carboxylesterase isozymes involved in metabolic activation of prodrugs. Molecules 2008;13:412–31.

[36] Lehmann H, Ryan E. The familial incidence of low pseudocholinesterase level. Lancet 1956;271:124.

[37] Jensen FS, Schwartz M, Viby-Mogensen J. Identification of human plasma cholinesterase variants using molecular biological techniques. Acta Anaesthesiol Scand 1995;39:142–9.

[38] Kalow W, Genest K. A method for the detection of atypical forms of human serum cholinesterase; determination of dibucaine numbers. Can J Biochem 1957;35:339–46.

[39] Kalow W, Staron N. On distribution and inheritance of atypical forms of human serum cholinesterase, as indicated by dibucaine numbers. Can J Biochem Physiol 1957;35:1305–20.

[40] Kalow W. Pharmacogenetics and anesthesia. Anesthesiology 1964;25:377–87.

[41] Stockley IH. Neuromuscular blocker and anaesthetic drug interactions. In: Stockley IH, editor. Drug interactions. London: Pharmaceutical Press; 1998. p. 722–56.

[42] Borlak J, Reamon-Buettner SM. N-acetyltransferase 2 (NAT2) gene polymorphisms in colon and lung cancer patients. BMC Medical Genetics 2006;7:58.

[43] Evans DA, Manley KA, McKusick VA. Genetic control of isoniazid metabolism in man. Br Med J 1960;2(5197):485–91.

[44] Butcher NJ, Boukouvala S, Sim E, Minchin RF. Pharmacogenetics of the arylamine N-acetyltransferases. Pharmacogenomics J 2002;2:30–42.

[45] Okkels H, Sigsgaard T, Wolf H, Autrup H. Arylamine N-acetyltransferase 1 (NAT1) and 2 (NAT2) polymorphisms in susceptibility to bladder cancer: the influence of smoking. Cancer Epidemiol Biomarkers Prev 1997;6:225–31.

[46] Tukey RH, Strassburg CP. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. Annu Rev Pharmacol Toxicol 2000;40:581–616.

[47] Wormhoudt LW, Commandeur JNM, Vermeulen NPE. Genetic polymorphisms of human N-acetyltransferase, cytochrome P450, glutathione-S-transferase, and epoxide hydrolase enzymes: relevance to xenobiotic metabolism and toxicity. Crit Rev Toxicol 1999;29(1):59–124.

[48] Bhasker CR, McKinnon W, Stone A, Lo AC, Kubota T, Ishizaki T, et al. Genetic polymorphism of UDP-glucuronosyltransferase 2B7 (UGT2B7) at amino acid 268: ethnic diversity of alleles and potential clinical significance. Pharmacogenetics 2000;10:679–85.

[49] Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Bélanger A, et al. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics 1997;7:255–69.

[50] Rollason V, Samer C, Piguet V, Dayer P, Desmeules J. Pharmacogenetics of analgesics: toward the individualization of prescription. Pharmacogenomics 2008;9:905–33.

[51] Chou WY, Wang CH, Liu PH, Liu CC, Tseng CC, Jawan B. Human opioid receptor A118G polymorphism affects intravenous patient-controlled analgesia morphine consumption after total abdominal hysterectomy. Anesthesiology 2006;105:334–7.

[52] Hayashida M, Nagashima M, Satoh Y, Katoh R, Tagami M, Ide S, et al. Analgesic requirements after major abdominal surgery are associated with OPRM1 gene polymorphism genotype and haplotype. Pharmacogenomics 2008;9:1605–16.

[53] Janicki PK, Schuler G, Francis D, Bohr A, Gordin V, Jarzembowski T, et al. A genetic association study of the functional A118G polymorphism of the human mu-opioid receptor gene in patients with acute and chronic pain. Anesth Analg 2006;103:1011–7.

[54] Nackley AG, Tan KS, Fecho K, Flood P, Diatchenko L, Maixner W. Catechol-O-methyltransferase inhibition increases pain sensitivity through activation of both beta(2)- and beta(3)-adrenergic receptors. Pain 2007;128:199–208.

[55] Oertel BG, Schmidt R, Schneider A, Geisslinger G, Lotsch J. The mu-opioid receptor gene polymorphism 118A.G depletes alfentanil-induced analgesia and protects against respiratory depression in homozygous carriers. Pharmacogenet Genomics 2006;16:625–36.

[56] Xie HG, Dishy V, Sofowora G, Kim RB, Landau R, Smiley RM, et al. Arg389Gly β_1 adrenoceptor polymorphism varies in frequency among different ethnic groups but does not alter response *in vivo*. Pharmacogenetics 2001;11:191–7.

[57] Puy H, Gouya L, Deybach JC. Porphyrias. Lancet 2010;375:924–37.

[58] Moore MR, McColl KM, Rimington C, Goldberg A. Disorders of porphyrin metabolism. Plenum Publishing; 1987.

[59] Anderson KE, Bloomer JR, Bonkovsky HL, Kushner JP, Pierach CA, Pimstone NR, et al. Recommendations for the diagnosis and treatment of the acute porphyrias. Ann Intern Med 2005;142:439–50.

[60] Kauppinen R, Mustajoki P. Prognosis of acute porphyria: occurrence of acute attacks, precipitating factors, and associated diseases. Medicine (Baltimore) 1992;71:1–13.

[61] Mustajoki P, Nordmann Y. Early administration of heme arginate for acute porphyric attacks. Arch Intern Med 1993;153:2004–8.

[62] Deybach JC, Badminton M, Puy H, Sandberg S, Frank J, Harper P, et al. European porphyria initiative (EPI): a platform to develop a common approach to the management of porphyrias and to promote research in the field. Physiol Res 2006;55(Suppl. 2):67–73.

[63] La Du BN. Butyrylcholinesterase variants and the new methods of molecular biology. Acta Anaesthesiol Scand 1995;39:139–41.

[64] Hopkins PM. Malignant hyperthermia: advances in clinical management and diagnosis. Br J Anaesth 2000;85:118–28.

[65] Ording H. Incidence of malignant hyperthermia in Denmark. Anesth Analg 1985;64:700–40.

[66] European Malignant Hyperthermia Group. <www.emhg.org>.

[67] Robinson RL, Hopkins PM, Carsana A, Gilly H, Halsall J, Heytens L, et al. Several interacting genes influence the malignant hyperthermia phenotype. Hum Genet 2003;112:217–8.

[68] Malignant Hyperthermia Association of the United States. 2013 <www.mhaus.org>.

[69] Kaufmann A, Kraft B, Michalek-Sauberer A, Weigl LG. Novel ryanodine receptor mutation that may cause malignant hyperthermia. Anesthesiology 2008;109:457–64.

[70] Levitt RC, Olckers A, Meyers S, Fletcher JE, Rosenberg H, Isaacs H, et al. Evidence for the localization of a malignant hyperthermia susceptibility locus (MHS2) to human chromosome 17q. Genomics 1992;14:562–6.

[71] Monnier N, Procaccio V, Stieglitz P, Lunardi J. Malignant hyperthermia susceptibility is associated with a mutation of the alpha 1-subunit of the human dihydropyridine-sensitive L-type voltage-dependent calcium-channel receptor in skeletal muscle. Am J Hum Genet 1997;60:1316–25.

[72] Monnier N, Kozak-Ribbens G, Krivosic-Horber R, Nivoche Y, Qi D, Kraev N, et al. Presence of two different genetic traits in malignant hyperthermia families: implications for genetic analysis, diagnosis, and incidence of malignant hyperthermia susceptibility. Anesthesiology 2002;97:1067–74.

[73] Denborough MA, Forster JF, Lovell RR, Maplestone PA, Villiers JD. Anaesthetic deaths in a family. Br J Anaesth 1962;34:395–6.

[74] Robinson RL, Carpenter D, Shaw M-A, Halsall PJ, Hopkins P. Mutations in RYR1 in malignant hyperthermia and central core disease. Hum Mutat 2006;27:977–89.

[75] Fletcher JE, Rosenberg H, Aggarwal M. Comparison of European and North American malignant hyperthermia diagnostic protocol outcomes for use in genetic studies. Anesthesiology 1999;90:654–61.

[76] Stamer UM, Stuber F. Pharmacogenomics and perioperative medicine: Implications for modern clinical practice. Can J Anaesth 2009;56:467–8.

[77] Urwyler A, Deufel T, McCarthy T, West S. Guidelines for the molecular detection of susceptibility to malignant hyperthermia. Br J Anaesth 2001;86:283–7.

[78] Kosuge K, Jun Y, Watanabe H, Kimura M, Nishimoto M, Ishizaki T, et al. Effects of CYP3A4 inhibition by diltiazem on pharmacokinetics and dynamics of diazepam in relation to CYP2C19 genotype status. Drug Metab Dispos 2001;29:1284–9.

[79] Qin XP, Xie HG, Wang W, He N, Huang SL, Xu ZH, et al. Effect of the gene dosage of CYP2C19 on diazepam metabolism in Chinese subjects. Clin Pharmacol Ther 1999;66:642–6.

[80] Shih PS, Huang JD. Pharmacokinetics of midazolam and 1′-hydroxymidazolam in Chinese with different CYP3A5 genotypes. Drug Metab Dispos 2002;30:1491–6.

[81] Liem EB, Lin CM, Suleman MI, Doufas AG, Gregg RG, Veauthier JM, et al. Anesthetic requirement is increased in redheads. Anesthesiology 2004;101:279–83.

[82] Ray DC, Drummond GB. Halothane hepatitis. Br J Anaesth 1991;67:84–99.

[83] Bruun LS, Elkjaer S, Bitsch-Larsen D, Andersen O. Hepatic failure in a child after acetaminophen and sevoflurane exposure. Anesth Analg 2001;92:1446–8.
[84] Nishiyama T, Yokoyama T, Hanaoka K. Liver function after sevoflurane or isoflurane anaesthesia in neurosurgical patients. Can J Anaesth 1998;45:753–6.
[85] Ohmori H, Seki S, Kanaya N, Imasaki H, Namaki A. A case report of postoperative liver dysfunction following sevoflurane anesthesia after isoflurane anesthesia. J Jpn Soc Clin Anesth 1994;14:68–71.
[86] Cerf C, Mesguish M, Gabriel I, Amselem S, Duvaldestin P. Screening patients with prolonged neuromuscular blockade after succinylcholine and mivacurium. Anesth Analg 2002;94:461–6.
[87] Ostergaard D, Jensen FS, Skovgaard LT, Viby-Mogensen J. Dose-response relationship for mivacurium in patients with phenotypically abnormal plasma cholinesterase activity. Acta Anaesthesiol Scand 1995;39:1016–8.
[88] Ikeda K, Ide S, Han WH, Hayashida M, Uhl GR, Sora I. How individual sensitivity to opiates can be predicted by gene analyses. Trends Pharmacol Sci 2005;26:311–7.
[89] Klepstad P, Rakvåg TT, Kaasa S, Holthe M, Dale O, Borchgrevink PC, et al. The 118 A.G polymorphism in the human mu-opioid receptor gene may increase morphine requirements in patients with pain caused by malignant disease. Acta Anaesthesiol Scand 2004;48:1232–9.
[90] Landau R, Kern C, Columb MO, Smiley RM, Blouin JL. Genetic variability of the mu-opioid receptor influences intrathecal fentanyl analgesia requirements in laboring women. Pain 2008;139:5–14.
[91] Somogyi AA, Barratt DT, Coller JK. Pharmacogenetics of opioids. Clin Pharmacol Ther 2007;81:429–44.
[92] Lovlie R, Daly AK, Matre GE, Molven A, Steen VM. Polymorphisms in CYP2D6 duplication-negative individuals with the ultrarapid metabolizer phenotype: a role for the CYP2D6*35 allele in ultrarapid metabolism? Pharmacogenetics 2001;11:45–55.
[93] Chou WY, Yang LC, Lu HF, Ko JY, Wang CH, Lin SH, et al. Association of mu-opioid receptor gene polymorphism (A118G) with variations in morphine consumption for analgesia after total knee arthroplasty. Acta Anaesthesiol Scand 2006;50:787–92.
[94] Meineke I, Freudenthaler S, Hofmann U, Schaeffeler E, Mikus G, Schwab M, et al. Pharmacokinetic modelling of morphine, morphine-3-glucuronide and morphine-6- glucuronide in plasma and cerebrospinal fluid of neurosurgical patients after short-term infusion of morphine. Br J Clin Pharmacol 2002;54:592–603.
[95] Chen ZR, Somogyi AA, Bochner F. Polymorphic O-demethylation of codeine. Lancet 1988;2:914–5.
[96] Chen ZR, Somogyi AA, Reynolds G, Bochner F. Disposition and metabolism of codeine after single and chronic doses in one poor and seven extensive metabolizers. Br J Clin Pharmacol 1991;31:381–90.
[97] Yue QY, Svensson JO, Alm C, Sjoqvist F, Sawe J. Codeine O-demethylation co-segregates with polymorphic debrisoquine hydroxylation. Br J Clin Pharmacol 1989;28:639–45.
[98] Kirchheiner J, Meineke I, Freytag G, Meisel C, Roots I, Brockmoller J. Enantiospecific effects of cytochrome P450 2C9 amino acid variants on ibuprofen pharmacokinetics and on the inhibition of cyclooxygenases 1 and 2. Clin Pharmacol Ther 2002;72:62–75.
[99] Williams DG, Patel A, Howard RF. Pharmacogenetics of codeine metabolism in an urban population of children and its implications for analgesic reliability. Br J Anaesth 2002;89:839–45.
[100] Gasche Y, Daali Y, Fathi M, Chiappe A, Cottini S, Dayer P, et al. Codeine intoxication associated with ultrarapid CYP2D6 metabolism. N Engl J Med 2004;351(27):2827–31.
[101] Eap CB, Broly F, Mino A, Hämmig R, Déglon JJ, Uehlinger C, et al. Cytochrome P450 2D6 genotype and methadone steady-state concentrations. J Clin Psychopharmacol 2001;21:229–34.
[102] Coller JK, Barratt DT, Dahlen K, Loennechen MH, Somogyi AA. ABCB1 genetic variability and methadone dosage requirements in opioid-dependent individuals. Clin Pharmacol Ther 2006;80:682–90.
[103] Enggaard TP, Poulsen L, Arendt-Nielsen L, Brosen K, Ossig J, Sindrup SH. The analgesic effect of tramadol after intravenous injection in healthy volunteers in relation to CYP2D6. Anesth Analg 2006;102:146–50.
[104] Poulsen L, Arendt-Nielsen L, Brosen K, Sindrup SH. The hypoalgesic effect of tramadol in relation to CYP2D6. Clin Pharmacol Ther 1996;60:636–44.
[105] Stamer UM, Musshoff F, Kobilay M, Madea B, Hoeft A, Stuber F. Concentrations of tramadol and O-desmethyltramadol enantiomers in different CYP2D6 genotypes. Clin Pharmacol Ther 2007;82:41–7.
[106] Kirchheiner J, Stormer E, Meisel C, Steinbach N, Roots I, Brockmoller J. Influence of CYP2C9 genetic polymorphisms on pharmacokinetics of celecoxib and its metabolites. Pharmacogenetics 2003;13:473–80.
[107] Sandberg M, Yasar U, Stromberg P, Hoog JO, Eliasson E. Oxidation of celecoxib by polymorphic cytochrome P450 2C9 and alcohol dehydrogenase. Br J Clin Pharmacol 2002;54:423–9.
[108] Bernal W, Donaldson P, Underhill J, Wendon J, Williams R. Tumor necrosis factor genomic polymorphism and outcome of acetaminophen (paracetamol)-induced acute liver failure. J Hepatol 1998;29:53–9.
[109] Rueffert H, Thieme V, Wallenborn J, Lemnitz N, Bergmann A, Rudlof K, et al. Do variations in the 5-HT3A and 5HT3B serotonin receptor genes (HTR3A, HTR3B) influence the occurrence of postoperative nausea and vomiting (PONV)? Anesth Analg 2009;109:1442–7.
[110] Mancinelli L, Cronin M, Sadée W. Pharmacogenomics: the promise of personalized medicine. AAPS Pharm Sci 2000;2:E4.

This page is intentionally left blank

Chapter 36

Tuberculosis

Geetha Ramachandran and Soumya Swaminathan
National Institute for Research in Tuberculosis, Indian Council of Medical Research, India

Chapter Outline

36.1 Overview 835
36.2 TB Etiology and Natural History 835
36.3 Burden of TB 836
  36.3.1 Incidence and Prevalence 836
  36.3.2 Mortality 838
36.4 Burden of TB Among Children 838
  36.4.1 Incidence, Mortality, and Prevalence 838
36.5 TB Diagnosis 838
  36.5.1 Active Disease and Latent TB Infection 838
  36.5.2 Extrapulmonary TB (EPTB) 839
  36.5.3 Diagnostic Methods: Serodiagnostic Tests, and Imaging Studies 840
36.6 Chemotherapy for TB 840
  36.6.1 Scientific Basis for Short-Course Chemotherapy 841
  36.6.2 Treatment of Smear-Positive TB 842
  36.6.3 MDR and XDR TB 842
36.7 Risk Factors: Genetics, HIV, Diabetes, and Smoking 842
36.8 TB in Children 844
  36.8.1 Development and Natural History 844
  36.8.2 Diagnosis and Treatment 844
36.9 Pharmacogenomics of TB 845
  36.9.1 Genetic Control of Drug Action 845
36.10 Pharmacogenomics and TB 845
  36.10.1 Isoniazid 845
  36.10.2 Rifampicin 852
  36.10.3 Aminoglycosides 853
  36.10.4 Predicting Drug-Induced Liver Injury 853
36.11 Conclusion 853
References 853

36.1 OVERVIEW

Tuberculosis (TB) remains a serious infectious disease around the world [1]. It causes ill health among millions of people each year and ranks as the second leading cause of death from an infectious disease worldwide, after the human immunodeficiency virus (HIV) [2]. This is despite the availability of treatment that will cure most TB cases.

The World Health Organization (WHO) declared TB a global public health emergency in 1993. Starting in the mid-1990s, efforts to improve TB care and control intensified at national and international levels. WHO developed the Directly Observed Treatment Short-Course (DOTS) strategy, a five-component package comprising political commitment, diagnosis using sputum smear microscopy, a regular supply of first-line anti-TB drugs, short-course chemotherapy, and a standard system for recording and reporting the number of cases detected.

36.2 TB ETIOLOGY AND NATURAL HISTORY

Tuberculosis is caused by the bacillus *Mycobacterium tuberculosis* (Figure 36.1) [3]. It typically affects the lungs (pulmonary TB) but can affect other sites as well (extrapulmonary TB). The disease is spread in the air when people who are sick with pulmonary TB expel bacteria, for example by coughing. In general, a relatively small proportion of people infected with *Mycobacterium tuberculosis* will develop TB disease; however, the probability of developing TB is much higher among people infected with HIV. TB is also more common among men than women and affects mostly adults in economically productive age groups.

The natural history of TB is influenced by several factors, the course being determined by the balance between host immunity and virulence of the TB bacillus. This understanding facilitates identification of areas where interventional strategies can be identified for TB control [4–6].

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00036-0
© 2014 Elsevier Inc. All rights reserved.

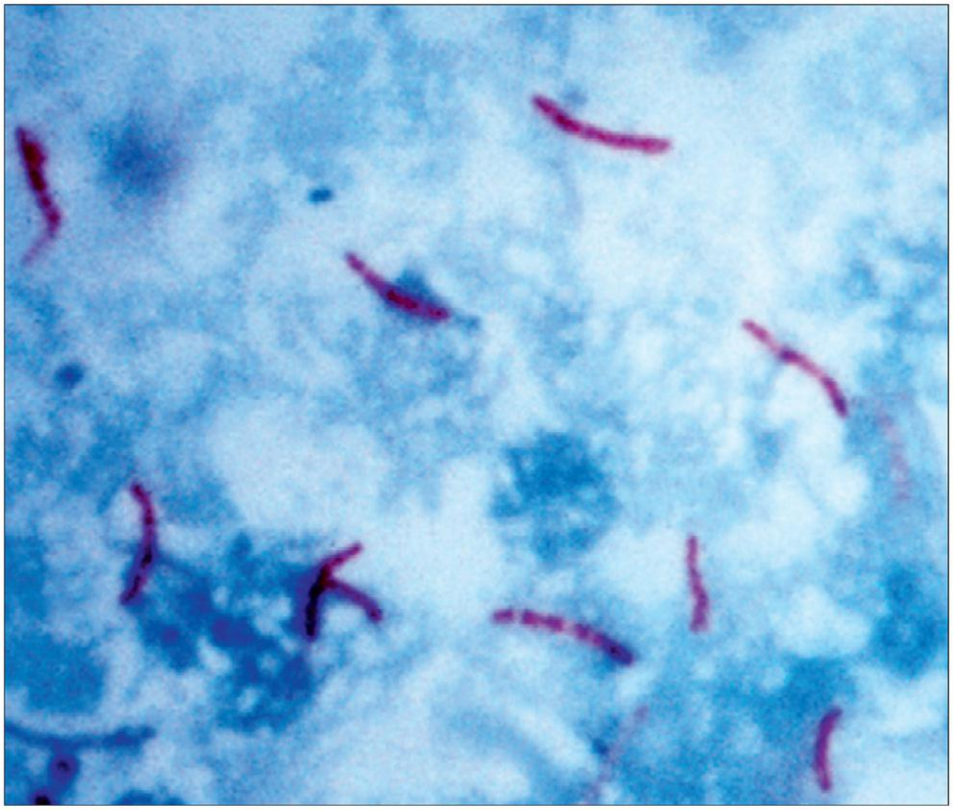


FIGURE 36.1 **Causative organism for TB.** High-power micrograph of acid-fast bacilli in the sputum of a patient with tuberculosis, shown by Ziehl-Neelsen staining (×1000). *Source: Reproduced with permission from Lawn and Zumla* [3].

36.3 BURDEN OF TB

The burden of disease caused by TB can be measured in terms of incidence (defined as the number of new and relapse cases arising in a given time period, usually one year), prevalence (defined as the number of cases at a given point in time), and mortality (defined as the number of deaths in a given time period, usually one year).

There has been major progress in reducing TB cases and deaths in the past two decades, with incidence having fallen globally for several years and declining at a rate of 2.2% between 2010 and 2011. Globally, the TB mortality rate has fallen by 41% since 1990 and the world is on track to reach the global target of a 50% reduction by 2015. Despite this encouraging progress, the global burden of TB remains enormous. There were an estimated 8.7 million cases of TB in 2011 (13% coinfected with HIV). There were also 1.4 million deaths from it (990,000 among HIV-negative individuals and 430,000 among HIV-positive individuals) (Figure 36.2) [7]. These deaths included 0.5 million among women, making TB one of the top killers of women worldwide [2].

Geographically, the burden of TB is highest in Asia and Africa. India and China combined have almost 40% of the world's TB cases; the Southeast Asian and Western Pacific regions, of which India and China are a part, account for 60%. The African region has approximately one-quarter of the world's cases, as well as the highest rates of cases and deaths relative to population. Globally, 3.7% of new and 20% of previously treated cases are estimated to have multidrug-resistant TB (MDR-TB). India, China, the Russian Federation, and South Africa have almost 60% of the world's cases. The highest proportions of TB patients with MDR-TB are in Eastern Europe and Central Asia. Almost 80% of TB cases among people living with HIV reside in Africa. There were an estimated 0.5 million cases and 64,000 deaths among children in 2011 [2].

36.3.1 Incidence and Prevalence

In 2011, there were an estimated 8.7 million incident cases of TB (range 8.3–9.0 million) globally, equivalent to 125 cases per 100,000 population. Most of these cases occurred in Asia (59%) and Africa (26%); smaller proportions of cases occurred in the Eastern Mediterranean region (7.7%), the European region (4.3%), and the Americas (3%). Of the 8.7 million incident cases, an estimated 0.5 million were children and 2.9 million (range 2.6–3.2 million) were women [2].

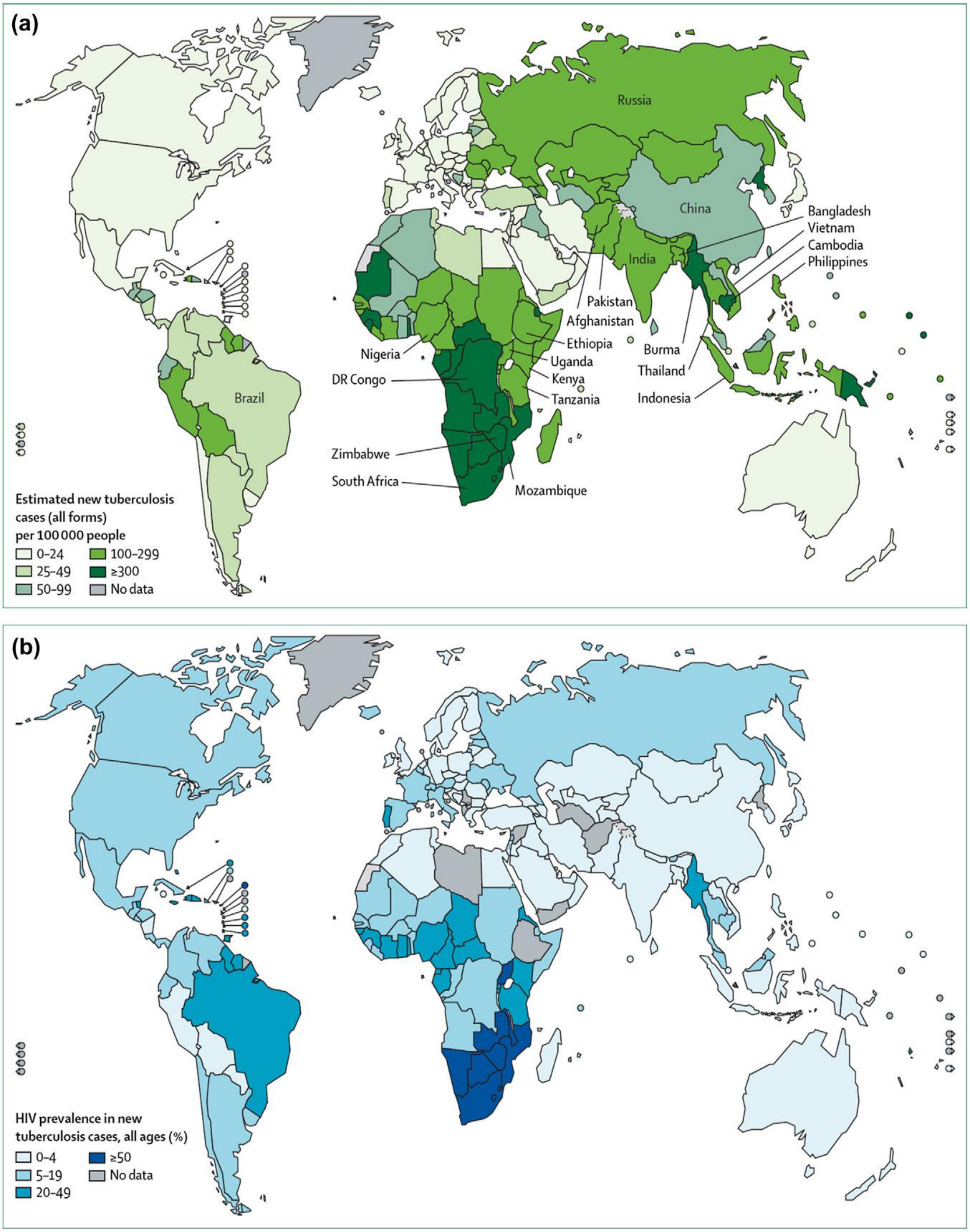

(a)
Russia
China
India
Brazil
Bangladesh
Vietnam
Cambodia
Philippines
Pakistan
Afghanistan
Ethiopia
Uganda
Kenya
Tanzania
Burma
Thailand
Indonesia
Nigeria
DR Congo
Zimbabwe
South Africa
Mozambique
Estimated new tuberculosis cases (all forms) per 100 000 people
0–24
25–49
50–99
100–299
≥300
No data
(b)
HIV prevalence in new tuberculosis cases, all ages (%)
0–4
5–19
20–49
≥50
No data

FIGURE 36.2 **TB incidence and HIV prevalence in new cases of TB.** (a) Estimated tuberculosis incidence (2010); data are from a WHO report. The 22 countries with the highest burden of tuberculosis—which account for 80% of cases worldwide—are labeled. (b) Estimated HIV prevalence in new cases of tuberculosis (2010). *Source: Reproduced with permission from Raviglione et al.* [7].

The five countries with the largest number of incident cases in 2011 were India (2.0–2.5 million), China (0.9–1.1 million), South Africa (0.4–0.6 million), Indonesia (0.4–0.5 million), and Pakistan (0.3–0.5 million). India and China alone accounted for 26% and 12% of global cases, respectively. Of the 8.7 million incident cases in 2011, 1.0–1.2 million (12–14%) were among people with HIV, with a best estimate of 1.1 million (13%). The proportion of TB cases coinfected with HIV was highest in countries in the African region (39%) which accounted for 79% of TB cases among people living with HIV worldwide [2].

Globally, incidence rates were relatively stable from 1990 up to around 2001 and then started to fall. Between 2010 and 2011, the rate of decline was 2.2%. The rate of decline between 2010 and 2011 was 0.5% in the Eastern Mediterranean region, 2.0% in the Southeast Asia region, 2.3% in the Western Pacific region, 3.1% in the African region, 3.8% in the Americas, and 8.5% in the European region. Incidence rates have been falling since the mid-1990s in the Eastern Mediterranean region and since around 2000 in Southeast Asia; they peaked at the end of the 1990s in the European region and around 2002 in Africa; they have been falling since 1990 in the Americas and the Western Pacific region [2].

There were an estimated 12 million prevalent TB cases (range 10 million–13 million) in 2011, equivalent to 170 cases per 100,000 population. Prevalence rates have fallen by 36% globally since 1990, and they are declining regionally. America halved its 1990 TB prevalence level by around 2005, and the Western Pacific region is close to doing so [2].

36.3.2 Mortality

The number of TB deaths worldwide fell just below 1 million among HIV-negative individuals in 2011. The number of TB deaths per 100,000 population in this group, plus the estimated number of TB deaths among HIV-positive individuals, equates to a best estimate of 20 deaths per 100,000 population in 2011. Globally, mortality rates (excluding deaths among HIV-positive individuals) have fallen by 41% since 1990. The current forecast suggests that a 50% reduction by 2015 compared with a baseline of 1990 will be achieved.

36.4 BURDEN OF TB AMONG CHILDREN

For many years, prevention, diagnosis, and treatment of TB in children have been relatively neglected. The greatest attention has been given to the detection and treatment of infectious cases, most of which occur in adults. To help address the burden of TB in children (defined as those aged <15 years) and monitor progress, robust data are necessary. WHO report (2012) on global TB care and control has included, for the first time estimates of the burden of TB among children, with best estimates of 490,000 cases and 64,000 deaths per year.

36.4.1 Incidence, Mortality, and Prevalence

There is no easy-to-use, accurate diagnostic test for TB in children. Most children have paucibacillary TB, which is harder to diagnose with sputum smear microscopy and culture. Many children, especially younger ones, are also not able to expectorate sputum. Diagnosis is usually made via a combination of clinical (as opposed to laboratory) criteria and a nonspecific test for tuberculous infection, but there is no universally applied diagnostic algorithm.

The definitive diagnosis of extrapulmonary TB requires specialized services that are usually available only in referral hospitals and thus not often accessible to those in need. Besides diagnostic challenges, children diagnosed with TB are not always reported to national surveillance systems because of the lack of linkages among individual pediatricians, pediatric hospitals, and national TB programs; also, data from national surveys including children are limited.

TB incidence among children was estimated at 490,000 (range 470,000–510,000) in 2011, which is equivalent to about 6% of the total number of 8.7 million incident cases [2]. The total number of deaths from TB among HIV-negative children was estimated at 64,000 (range 58,000–71,000) in 2011, which is equivalent to 6% of the 990,000 TB deaths among HIV-negative TB cases in 2011 [2]. Besides the direct impact of TB on children themselves, parental TB deaths have created large numbers of orphans. In 2009, there were almost 10 million children who were orphans as a consequence of losing at least one parent to TB.

Data on the prevalence of TB in children are limited to a few nationwide surveys conducted before 2001. Examples include a survey in India in 1956 and surveys in China in 1980, 1990, and 2000. A 2007 survey in the Philippines included children aged 10–14 years. These surveys consistently found a low burden of bacteriologically confirmed TB in children compared with adults [2].

36.5 TB DIAGNOSIS

36.5.1 Active Disease and Latent TB Infection

Sputum smear microscopy, the primary method for diagnosis of pulmonary TB, is a simple, rapid, and inexpensive technique that is highly specific in areas with very high prevalence [8]. Its limitations include low sensitivity and diagnosis of extrapulmonary TB, pediatric TB, and patients coinfected with HIV [9]. Although fluorescent microscopy has improved the sensitivity of sputum smear microscopy, it involves major cost constraints [10]. These may be circumvented by the use of light-emitting diodes (LEDs), which are relatively cheaper [11].

Rapid culture-based methods for TB diagnosis include rapid automated liquid culture, thin-layer agar culture, phage-based culture, and microscopic observation drug susceptibility assay (MODS) [12]. However, while performance indicators on these techniques might be better, turnaround times are longer. There are also requirements in terms of investment in infrastructure and equipment [13]. Nucleic acid amplification tests (NAATs) attempt to provide accurate and rapid TB diagnosis using a technology that provides improved sensitivity and specificity functions as compared to sputum smear microscopy.

The Xpert MTB/RIF assay, which is based on molecular detection, enables simultaneous detection of *M. tuberculosis* and rifampicin resistance. It has been endorsed by WHO and is recommended as an initial diagnostic test for smear-negative patients and those at high risk for rifampicin (RMP) resistance. Although the development of the Xpert MTB/RIF assay is undoubtedly a landmark event, clinical and programmatic effects and cost-effectiveness remain to be defined [14]. A schematic of current TB diagnostic methods is shown in Figure 36.3 [14].

Diagnosis of latent TB infection (LTBI) is considered important as a tool for assessing the burden of TB for epidemiological purposes. Because LTBI contributes significantly to the pool of active TB cases later on, its early recognition is assuming importance in high-risk groups where there is a potential for treatment for this condition [15]. The tuberculin skin test (TST) and interferon-gamma release assays (IGRAs) are two diagnostic tests for the detection of LTBI [16]; both have advantages and limitations. In high-burden TB countries, WHO recommends that neither IGRAs nor TST be used to diagnose active TB and that IGRAs not be used to diagnose LTBI [17,18].

36.5.2 Extrapulmonary TB (EPTB)

The diagnostic yield of conventional AFB smear and mycobacterial culture varies with extrapulmonary fluid/specimen type [19,20]. A battery of diagnostic tests, including culture (and DST where feasible), enhances diagnostic yield in EPTB. PCR-based methods, adenosine deaminase, and interferon gamma are also used in EPTB diagnosis. There is

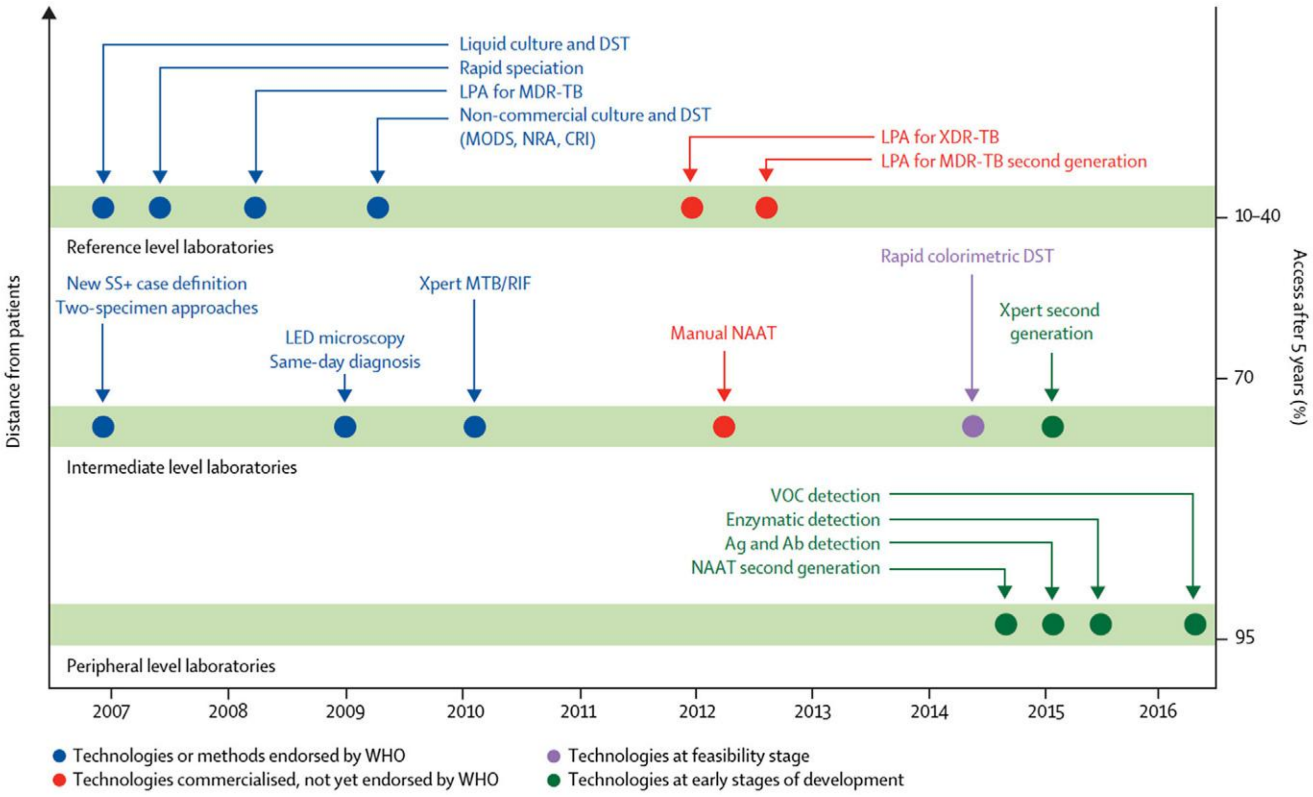

Liquid culture and DST
Rapid speciation
LPA for MDR-TB
Non-commercial culture and DST
(MODS, NRA, CRI)
LPA for XDR-TB
LPA for MDR-TB second generation
Reference level laboratories
New SS+ case definition
Two-specimen approaches
Xpert MTB/RIF
LED microscopy
Same-day diagnosis
Manual NAAT
Rapid colorimetric DST
Xpert second generation
Intermediate level laboratories
VOC detection
Enzymatic detection
Ag and Ab detection
NAAT second generation
Peripheral level laboratories
Distance from patients
Access after 5 years (%)
10–40
70
95
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
Technologies or methods endorsed by WHO
Technologies commercialised, not yet endorsed by WHO
Technologies at feasibility stage
Technologies at early stages of development

FIGURE 36.3 **Development pipeline for new tuberculosis diagnostics.** Reference-level laboratories refer to national-level facilities. Intermediate-level laboratories refer to district- and subdistrict-level facilities. Peripheral-level laboratories refer to community-level facilities. Ab—antibody; Ag—antigen; CRI—colorimetric redox indicator assay; DST—drug susceptibility test; LED—light-emitting diode; LPA—line-probe assay; MDR-TB—multidrug-resistant tuberculosis; MODS—microscopic observation drug susceptibility; NAAT—nucleic acid amplification tests; NRA—nitrate reductase assay; SS+—sputum smear–positive; VOC—volatile organic compound; XDR-TB—extensively drug-resistant tuberculosis. *Source: Reproduced from WHO's Global Tuberculosis Control report, 2012, by permission of the World Health Organization. Reproduced with permission from Lawn et al.* [14].

a need to systematically evaluate and establish the utility of newer tests such as GeneXpert in early diagnosis to increase diagnostic yield.

36.5.3 Diagnostic Methods: Serodiagnostic Tests, and Imaging Studies

A broad range of serodiagnostic tests with a varying degree of reliability, repeatability, and concordance have been marketed for TB diagnosis. Since no commercial serological assay consistently results in an accurate diagnosis. WHO recently recommended that they not be used for the diagnosis of pulmonary or extrapulmonary TB [21].

Imaging modalities, such as conventional X-ray, ultrasonography, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography CT (PET-CT) have been used to localize the disease, assess the extent of organ involvement, and evaluate treatment response. The chest X-ray is the mainstay of pulmonary TB imaging. However, it can be normal in HIV-infected patients with active TB disease as well as in some patients with miliary TB [22].

36.6 CHEMOTHERAPY FOR TB

The foundations of chemotherapy are (1) the prevention of the emergence of drug-resistant *M. tuberculosis*; (2) the bactericidal activity of drugs; and (3) the principles of intermittent drug dosage and mycobacteria lag phase.

Chemotherapy for TB became possible only after the discovery of streptomycin in 1944. The introduction of isoniazid (INH) as a drug in the combination regimen for TB treatment formed the basis of primary chemotherapy in the 1950s and 1960s. During that time, the standard regimen comprised streptomycin, INH, and PAS for a few months, followed by the latter two drugs for 18 months. Following the discovery of rifampicin in 1965, short-course chemotherapy for TB was introduced in the 1970s [23]. Figure 36.4 is a schematic representation of anti-TB drug development [24].

The aims of drug treatment of TB are

- To cure patients by the shortest duration of drug administration, preferably with minimum interference with their activities of daily living
- To prevent death from TB or late effects
- To prevent relapse
- To reduce transmission to the community

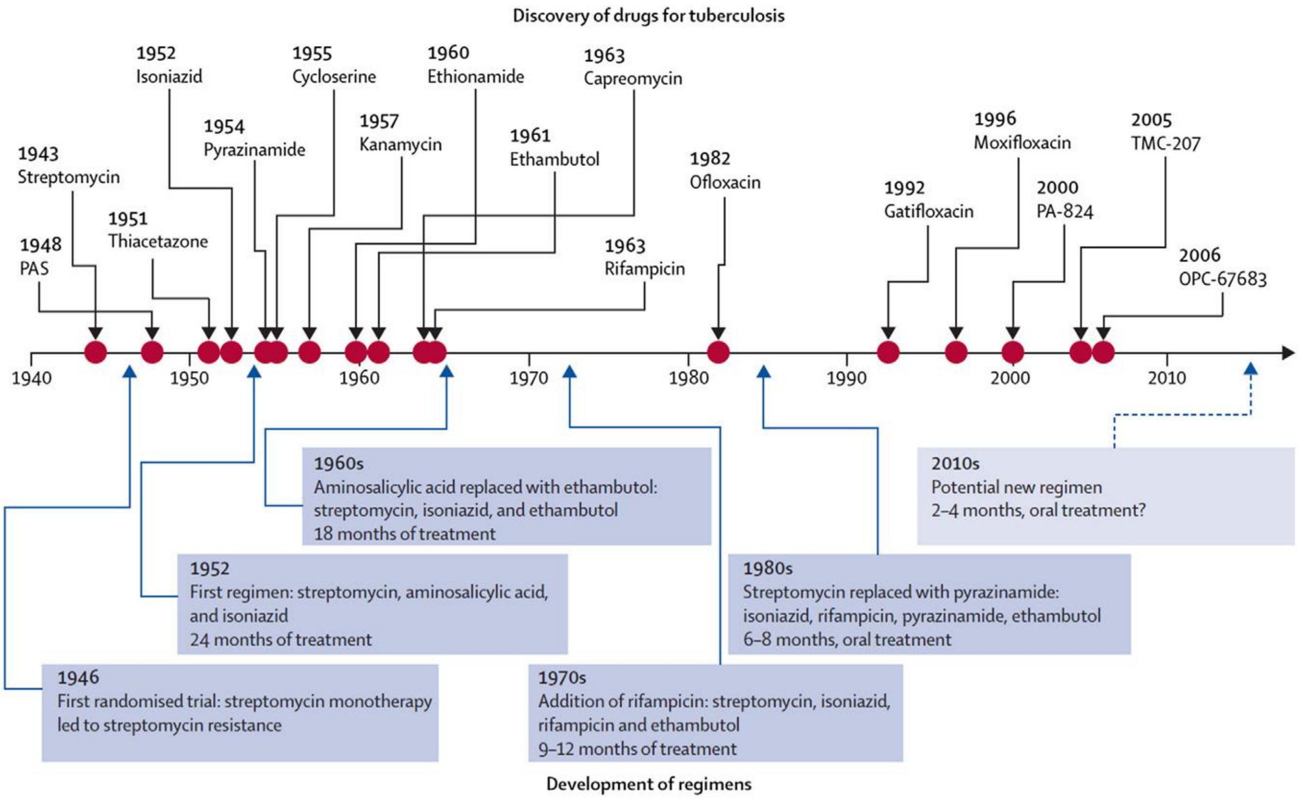

Discovery of drugs for tuberculosis
1943 Streptomycin
1948 PAS
1951 Thiacetazone
1952 Isoniazid
1954 Pyrazinamide
1955 Cycloserine
1957 Kanamycin
1960 Ethionamide
1961 Ethambutol
1963 Capreomycin
1963 Rifampicin
1982 Ofloxacin
1992 Gatifloxacin
1996 Moxifloxacin
2000 PA-824
2005 TMC-207
2006 OPC-67683
1940
1950
1960
1970
1980
1990
2000
2010
1946 First randomised trial: streptomycin monotherapy led to streptomycin resistance
1952 First regimen: streptomycin, aminosalicylic acid, and isoniazid 24 months of treatment
1960s Aminosalicylic acid replaced with ethambutol: streptomycin, isoniazid, and ethambutol 18 months of treatment
1970s Addition of rifampicin: streptomycin, isoniazid, rifampicin and ethambutol 9–12 months of treatment
1980s Streptomycin replaced with pyrazinamide: isoniazid, rifampicin, pyrazinamide, ethambutol 6–8 months, oral treatment
2010s Potential new regimen 2–4 months, oral treatment?
Development of regimens

FIGURE 36.4 **History of drug discovery and treatment regimens for tuberculosis compounds.** These compounds are in the early stage of development, but for which there are no human proof-of-concept data, are not shown. The dashed-line arrow represents a future regimen. Red dots represent when the drugs were first reported. *Source: Reproduced with permission from Ma et al.* [24].

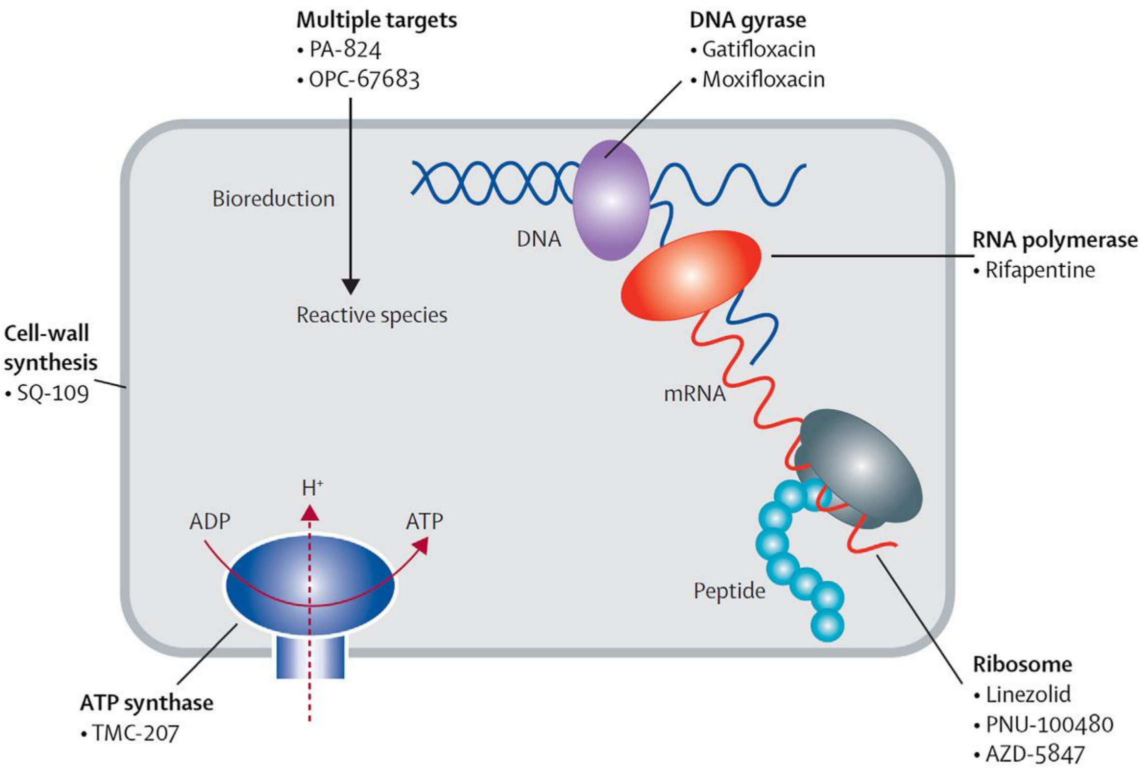


FIGURE 36.5 **Mechanisms of action for new compounds in clinical development for tuberculosis.** *Source: Reproduced with permission from Ma et al.* [24].

Figure 36.5 [24] shows the mechanism of action of anti-TB drugs currently being developed for treating TB. An important research priority is development of new drugs, given that no new drug has emerged since the discovery of RMP. The global pipeline for new anti-TB drugs is shown in Figure 36.6 [25].

36.6.1 Scientific Basis for Short-Course Chemotherapy

M. tuberculosis, the causative bacterium of TB, is slow-growing and can enter a phase of dormancy. A patient with TB can basically harbor four populations of organisms. The first is the actively growing extracellular organisms, which are present in abundance in aerated cavities. The second is slow intermittently growing organisms in an unstable part of the lesion. The third includes organisms surviving in a low environmental pH, which can occur in inflammatory lesions or in phagolysosomes or macrophages. The fourth is completely dormant organisms surviving under anaerobic conditions. The major actions of anti-TB drugs [26] are as follows:

- Bactericidal action, defined as the ability to rapidly kill actively growing bacilli
- Sterilizing action, defined as the capacity to kill semi-dormant organisms
- Prevention of resistance

INH is the most potent bactericidal drug followed by RMP. RMP and pyrazinamide (PZA) are capable of sterilizing tuberculous lesions and preventing disease relapse. Resistance to an anti-TB drug is due to spontaneous chromosomal mutation at a frequency of 10^{-6}–10^{-8} bacterial replications [27]. As mutations resulting in drug resistance are unlinked, the probability of resistance to all three drugs used simultaneously becomes 10^{-18}–10^{-20} [28]. Thus, the chance of drug resistance is practically nil when three or more effective drugs are used in combination. Among the first-line drugs, INH and RMP are the most effective in preventing resistance [26]; streptomycin (Sm), ethambutol (EMB), and PZA are less effective.

Drug-resistant TB usually results from poor patient adherence and other causes of failure to implement TB control programs. Poor treatment adherence can be multifaceted and complex, involving the interplay of individual patient characteristics and societal and economic factors.

Although the main aim in designing drug regimens has been prevention of drug resistance, more recently, increasing attention has been paid to methods of improving of

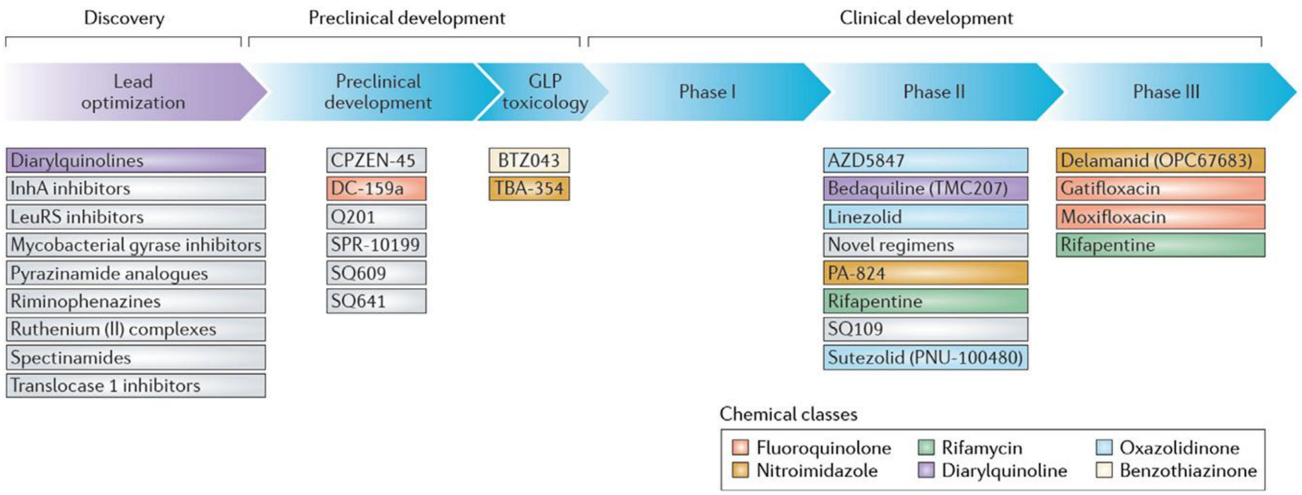

Discovery
Preclinical development
Clinical development
Lead optimization
Preclinical development
GLP toxicology
Phase I
Phase II
Phase III
Diarylquinolines
InhA inhibitors
LeuRS inhibitors
Mycobacterial gyrase inhibitors
Pyrazinamide analogues
Riminophenazines
Ruthenium (II) complexes
Spectinamides
Translocase 1 inhibitors
CPZEN-45
DC-159a
Q201
SPR-10199
SQ609
SQ641
BTZ043
TBA-354
AZD5847
Bedaquiline (TMC207)
Linezolid
Novel regimens
PA-824
Rifapentine
SQ109
Sutezolid (PNU-100480)
Delamanid (OPC67683)
Gatifloxacin
Moxifloxacin
Rifapentine
Chemical classes
Fluoroquinolone
Nitroimidazole
Rifamycin
Diarylquinoline
Oxazolidinone
Benzothiazinone

FIGURE 36.6 **Current global pipeline of new tuberculosis drugs.** Agents currently in discovery or development for the treatment of tuberculosis (TB) are shown. Ongoing projects without a lead compound series can be viewed at the website of the Working Group on New TB Drugs—Discovery Portfolio. GLP—good laboratory practice; InhA—enoyl-CoA reductase; LeuRS—leucyl-tRNA synthetase. *Source: Reproduced with permission from Zumla et al.* [25].

healthcare delivery and patient compliance. This can be accomplished by fully supervised drug administration.

36.6.2 Treatment of Smear-Positive TB

In the past few decades, a number of effective drug regimens have been found, largely through clinical trials for treating patients with newly diagnosed smear-positive pulmonary TB [29–32]. Most regimens are given for six months, this currently being the shortest duration of treatment required. A four-drug regimen (for two months) followed by two drugs (for four months) is advisable. Thus, the modern six-month short-course regimen comprises RMP, INH, PZA, and EMB for two months (intensive phase) followed by RMP and INH for four months (continuation phase), either daily or intermittently. For individual cases with extensive disease or cavitation on chest Xray and slow sputum bacteriological conversion to negative, the intensive phase may be prolonged beyond two months.

36.6.3 MDR and XDR TB

Drug resistance is a widespread phenomenon. MDR TB is defined as resistance to INH and RMP, the two key first-line anti-TB drugs. Extensively drug-resistant (XDR TB) is defined as resistance not only to RMP and INH but also to any fluoroquinolone and a second-line injectable drug [33]. Recent global epidemiological data on M/XDR-TB are shown in Figure 36.7 [34]. Although DOTS is highly effective in the management of drug-susceptible TB, it is not sufficient for controlling established MDR TB. Treatment success rates for patients with MDR TB are much lower compared with those for patients with drug- susceptible disease—namely 56–80% versus ≥90% [35–37]. On the basis of available data, WHO has recommended a three-part response to the global threat of MDR TB [37,38]—the widespread implementation of DOTS as the cornerstone of good TB control, improved drug susceptibility testing and surveillance, and rapid expansion of second-line drug therapy. This new initiative, known as DOTS-Plus [36] is currently coordinated by WHO in partnership with many agencies and institutions worldwide. The most important determinant in formulating a drug regimen for MDR TB is the patient's current drug susceptibility pattern *in vitro* [32], but most countries use standardized regimens.

A patient with MDR TB should receive a regimen comprising at least four to five second-line drugs for the initial months and three to four drugs subsequently. One regimen recommended by WHO consists of treatment with EMB, ethionamide/prothionamide, ofloxacin/ciprofloxacin, PZA, and aminoglycoside for 3 months, with the first four drugs being administered for at least a further 18 months [38,39].

36.7 RISK FACTORS: GENETICS, HIV, DIABETES, AND SMOKING

Several genetic, social, environmental, and biological determinants of health have been intuitively recognized by clinicians as risk factors for TB. Some of these are discussed in the following sections.

Certain key issues should be considered when evaluating genetic susceptibility to TB disease. Susceptibility does not follow a Mendelian pattern and is polygenic and multifactorial. The presence of two different genomes (of the TB bacillus and the host) and their interaction can influence the disease [40]. Several reports have associated certain genes with TB risk

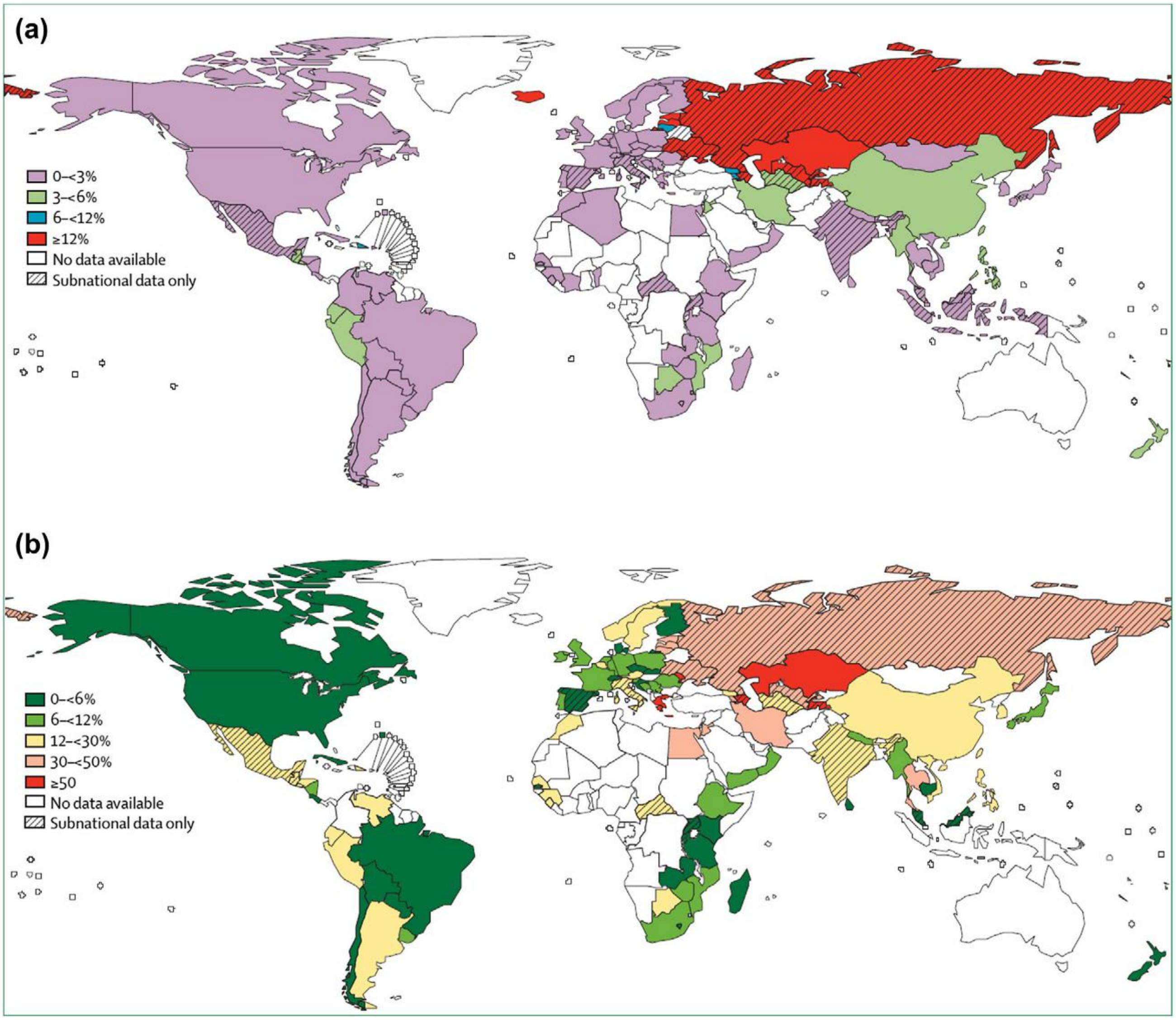


FIGURE 36.7 **Proportion of multidrug-resistant diseases.** Data shown are for (A) new cases and (B) previously treated cases of tuberculosis. *Source: Based on Global Drug Resistance Surveillance Project data. Australia, Democratic Republic of the Congo, Fiji, Guam, New Caledonia, Solomon Islands, and Qatar reported data for combined new and previously treated cases. Source: Reproduced with permission from Gandhi et al.* [34].

(including natural resistance–associated macrophage protein 1, interferon γ, nitric oxide synthase 2A, mannan-binding lectin, vitamin D receptor, and some Toll-like receptors) [41,42].

HIV infection and AIDS stand out as the most significant among all risk factors for TB and have consistently and significantly altered the incidence rate over the last three decades [43–46]. The impact of HIV/AIDS has been most profound on HIV prevalence in Sub-Saharan Africa, where a dramatic increase in TB notification rates has been documented concurrent with increasing HIV prevalence. Among persons living with HIV, TB can develop at any stage of infection and there is strong evidence that a declining CD4+ T-lymphocyte count and a high viral load are risk factors for disease. Treatment with highly active antiretroviral therapy (HAART) reduces risk [43].

Even though, several institutional outbreaks of MDR-TB among HIV-infected patients drew attention to the problem two decades ago [47–49], according to available evidence, HIV infection per se does not appear to be a risk factor for MDR-TB.

The lethal interaction between diabetes mellitus (DM) and TB is being increasingly recognized the world over [50,51]. Epidemiological modeling data suggest that, in India, 14.8% of all pulmonary TB cases and 20% of sputum smear–positive cases are attributable to DM [51], suggesting that DM substantially contributes to the burden of TB, especially sputum smear–positive pulmonary TB in India.

Data from recent systematic reviews on tobacco smoking and TB suggest that tobacco smokers have about a 3-fold higher risk of TB than do nonsmokers, even after adjustment for other factors [52,53].

36.8 TB IN CHILDREN

Tuberculosis in childhood is the consequence of droplet infection with *Mycobacterium tuberculosis* spread by aerosol formation from sputum expectorated mostly by adults with pulmonary TB. Because of the broad base of the population pyramid in developing communities, considerably more young children are infected and become diseased than is the case in a developed community. Also, because more severe forms of disease tend to occur in younger children, morbidity and mortality resulting from childhood TB are proportionately higher in the developing world.

36.8.1 Development and Natural History

Children in close household contact with an adult, often a parent, grandparent, or other adult family member, with "open," or sputum smear–positive, TB, experience considerably higher morbidity and mortality than do those in contact with adults whose sputum is positive on culture only [54]. However, only a minority of children develop disease following infection. The risk of disease varies in different communities, from approximately 10–30% and is concentrated in the 3–12 months following infection [55]. Risk is further influenced by the age of the child. The lifetime risk of developing disease after infection is 43% in infants, 24% between 1 and 5 years age, and 15% in adolescents, compared to immunocompetent adults, who have a lifetime risk of 5–10% [56]. This risk, as well as morbidity and mortality, is the highest in very young children. Young children are particularly susceptible to disseminated forms of TB such as miliary TB and tuberculous meningitis.

The natural history of TB in childhood is well documented from close observations made by experienced pediatricians in the pre-chemotherapy era [57]. Following infection, *M. tuberculosis* spreads from the primary focus and is soon found in the regional lymph nodes. Further lymphohematogenous spread distributes the organism throughout the body. Small foci of viable organisms are established in vulnerable sites, such as the meninges, the bone ends, the apices of the lungs (Simon foci), and the kidneys. This period of dissemination ends 6–8 weeks after infection, at which point the tuberculin test becomes positive, a nondescript "fever of initiation" may be experienced [58], and the primary focus and primary complex enlarge and may become visible on chest X-rays. A period of great danger for the development of disease follows; most disease will manifest itself in children in the next 3–24 months.

36.8.2 Diagnosis and Treatment

In only a minority of cases of childhood respiratory TB will a culture of *M. tuberculosis*, usually from an early morning gastric aspirate, be obtained. In most instances, a clinical diagnosis is made by drawing on the result of tuberculin testing, clinical features, a history of close contact with an adult with pulmonary TB, and chest X-ray. In developed countries, most cases of childhood TB are identified following contact tracing [59].

Induced sputum or early morning gastric aspirates are the most appropriate specimens in young children, but other specimens such as nasopharyngeal swabs or biopsy of lymph nodes or other tissues should not be ignored. Chest X-rays plays a major role in the diagnosis of childhood TB and the radiological features of childhood TB are well known.

The principle governing the treatment of pulmonary TB in children is similar to that followed in adults, which is a six-month short course regimen comprising of INH, RMP, EMB and PZA for two months (intensive phase) followed by INH and RMP for four months (continuation phase) [60]. The response to treatment of childhood respiratory TB is best judged by clinical features, in particular, weight gain. Abnormalities on chest X-ray are slow to respond, and mediastinal adenopathy may remain visible for up to 24 months, even following successful chemotherapy.

36.8.2.1 Drug-Resistant TB

Because of the relatively low mycobacterial burden in childhood TB, which is only intermittently metabolically active or dormant, acquired drug resistance in young childhood is very unusual. Depending on the incidence of drug resistance in a particular community, primary drug resistance in childhood is more common. INH-resistant and multidrug-resistant (MDR) *M. tuberculosis* strains had once been postulated to be less infectious and less likely to cause disease than their drug-susceptible counterparts, but studies have shown them to be just as likely to cause infection and disease [61].

Clinical and chest X-ray features are similar in children with drug-resistant and drug-susceptible TB, and so they do not help to identify resistant cases. Treatment of MDR TB cases in children need careful consideration of drug susceptibility patterns in the index case or, if available, the child's *M. tuberculosis* strain; the extent of the disease; and the possible toxicity of the anti-TB drugs considered (although toxicity is often less common in children). In general, a regimen of at least four anti-TB drugs to which the isolate is susceptible is employed. A fluoroquinolone and an injectable agent are most important for MDR cases; PZA, a sterilizing agent, may be continued because of the difficulty of performing a drug susceptibility test against this agent. Optimal duration of treatment is unknown, but 9–12 months after the last positive culture seems to be sufficient, depending on the severity of the disease. WHO recommends a standardized regimen for MDR TB that is followed in most developing countries. Clinical and chest X-ray follow-up after treatment completion is recommended.

36.9 PHARMACOGENOMICS OF TB

The science of pharmacogenomics—that is, defining the genetic determinants of drug effects—has evolved over the past 50 years. It aims to identify the inherited basis for interindividual differences in drug response and translate this into molecular diagnostics that can be used to individualize drug therapy. Individuals respond differently to drugs, and sometimes the effects are unpredictable. Differences in DNA sequences that alter the expression or function of drug-targeted proteins can contribute significantly to response variation in individuals.

The idea that genes control some drug responses was suggested in the 1950s because of the link between race or ethnicity and aberrant drug responses [62]. Subsequently, genes responsible for several inherited differences in drug metabolism and disposition were isolated, characterized, and linked to differences in drug effects [62]. Polymorphisms in genes encoding drug transporters and targets were also identified and shown to alter drug response [63,64].

It is well recognized that different patients respond in different ways to the same dose of a medication. These differences are often greater among members of a population than they are in the same person at different times [65]. The existence of large population differences with small intrapatient variability is consistent with inheritance as a determinant of drug response; it is estimated that genetics can account for 20–95% of variability in drug disposition and effects [66].

Although many nongenetic factors influence the effects of medications, including age, organ function, concomitant therapy, drug interactions, and nature of disease, there are now several examples of interindividual differences in drug response that are due to sequence variants in genes encoding drug-metabolizing enzymes, drug transporters, or drug targets [63]. Unlike other factors influencing drug response, inherited determinants generally remain stable throughout a lifetime [67].

More than 1.4 million single-nucleotide polymorphisms (SNPs) were identified in the initial sequencing of the human genome [67], with over 60,000 of them in the coding region of genes. Most drug effects are determined by the interplay of several gene products that influence the pharmacokinetics and pharmacodynamics of medications, including inherited differences in drug targets (e.g., receptors) and drug disposition (e.g., metabolizing enzymes and transporters).

36.9.1 Genetic Control of Drug Action

An individual's genetic endowment, phenotypically in protein structure, concentration, and configuration, may alter drug action in multiple ways. A drug entering the body comes into contact with numerous enzymes and other proteins. Nearly all drugs undergo enzymatically controlled transformations during their passage through the liver and other organs. Theoretically, genetic mutations that alter the quality of any of these proteins can lead to a recognizable disturbance in drug–cell interactions. For instance, variations in enzyme components of drug metabolism and conjugation may retard drug inactivation and consequent excretion, resulting in toxic concentrations after small doses. They may modify the site of enzymatic attack on the drug molecule, producing different metabolites with different properties.

Alternatively, genetic mutations may increase the activity or inducibility of drug-metabolizing enzymes and prevent the drug from reaching effective concentrations in plasma or tissues. Structural differences in serum proteins can presumably change binding affinities and alter the ratios of bound to free drug. Similarly, aberrant products at the site of a drug's action in organs, tissues, or cells may confer increased or decreased responsiveness to usual therapeutic drug concentrations.

Pharmacogenetic conditions may exist that result from the action of a single mutant gene. This leads to discontinuous variation in the measure of drug action or metabolism, and the population can be divided into two or more groups (with continuous variation in each) suggesting single-gene polymorphism. However, the variations most frequently observed have a Gaussian distribution resulting from polygenic and /or environmental factors or influences.

36.10 PHARMACOGENOMICS AND TB

Pharmacogenomics is the study of genetically determined, molecular bases of drug therapy outcomes. Outcomes can be broadly divided into areas of efficacy and toxicity.

TB is treated with a combination regimen that includes first-line drugs, INH, RMP, PZA, and EMB. Genetic factors are widely reported to play a significant role in modulating INH metabolism and drug/metabolite concentrations; this has a varying impact on treatment outcomes in terms of both efficacy and adverse events [68–71]. There are few reports that suggest gene polymorphisms as affecting RMP metabolism [72,73] and causing aminoglycoside-induced ototoxicity [74].

36.10.1 Isoniazid

Since its introduction in 1952, isoniazid (isonicotinyl hydrazine; INH) continues to be the most widely used chemotherapeutic agent for the treatment of TB [23]. It is highly efficacious and quite specific in its action against tubercle bacilli; also it has minimal toxicity and is less expensive, making it an important first-line anti-TB drug.

36.10.1.1 Metabolism

The pharmacology of INH has been extensively investigated [75]. Evidence from several studies [75,76] confirms that acetylation is the most important pathway for the metabolism of INH, and that difference in the rates of acetylation are determined by the INH acetylator status of individuals. INH is metabolized by acetylation, which takes place mainly in the liver and gut mucosa. This conjugation reaction is catalyzed by N-acetyl transferase (NAT2), during which an acetyl group from acetyl coenzyme A is transferred to the acceptor amine, thereby forming an amide. INH acetylation results in the formation of acetyl INH, which is conjugated with pyruvic acid and α ketoglutaric acid to form hydrazones; acetyl INH is hydrolyzed to monoacetyl hydrazine and isonicotinic acid. Monoacetyl hydrazine undergoes acetylation to form diacetyl hydrazine; isonicotinic acid is conjugated with glycine to form isonicotinyl glycine. A summary of INH metabolism is shown in Figure 36.8 [77].

36.10.1.2 Acetylator Genotypes and Phenotypes

Human acetylation rate is genetically determined; this is due to the polymorphic enzyme NAT, which is responsible for the metabolism of INH to acetyl INH. This enzyme activity is expressed at highly variable levels. Several studies over the years have concluded that human subjects show broad variations in their capacity to acetylate or inactivate INH to acetyl INH, the metabolites being devoid of anti-TB activity. Studies on TB patients treated for many months with INH-containing regimens have demonstrated that INH metabolism is noninducible.

Evidence has been obtained showing that rapid acetylators metabolize INH more rapidly than do slow acetylators; similar conclusions have been drawn from *in vitro* studies using liver biopsies from previously phenotyped individuals. Thus, the rate of INH acetylation varies from individual to individual because of differences in concentrations of the cytosolic NAT2 enzyme in liver and gut mucosa.

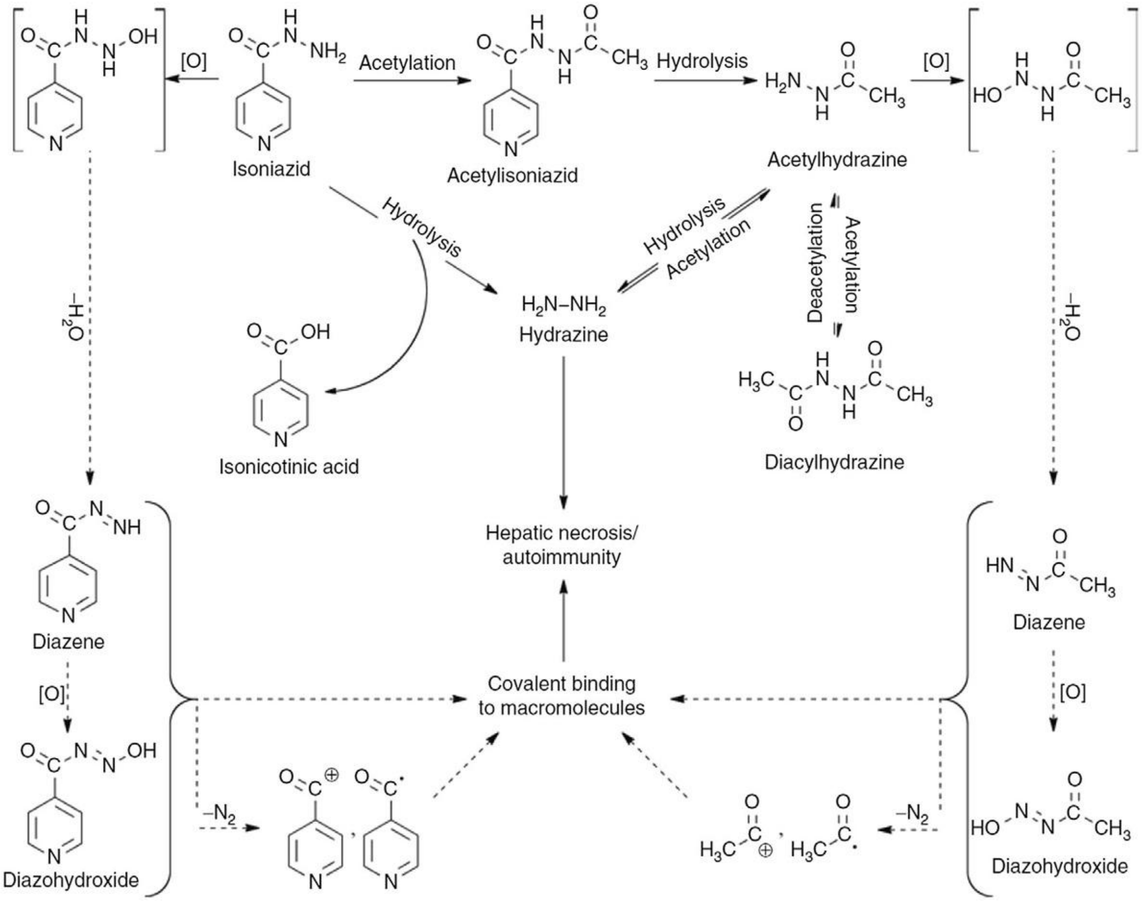

[O]
Isoniazid
Acetylation
Acetylisoniazid
Hydrolysis
Acetylhydrazine
[O]
Hydrolysis
Hydrolysis
Acetylation
Deacetylation
Acetylation
$H_2N{-}NH_2$
Hydrazine
Isonicotinic acid
Diacylhydrazine
$-H_2O$
$-H_2O$
Hepatic necrosis/
autoimmunity
Diazene
Diazene
[O]
[O]
Covalent binding
to macromolecules
$-N_2$
$-N_2$
Diazohydroxide
Diazohydroxide

FIGURE 36.8 **Proposed isoniazid metabolic pathways.** Solid lines represent major metabolic pathways; dashed lines represent speculated pathways. [O]—oxidation. *Source: Reproduced with permission from Metushi et al.* [77].

The difference in rapid and slow acetylators is due to the difference in the quantity rather than the quality of the enzyme, rapid acetylators having four to five times the quantity of the enzyme as the slow acetylators do. Variability in *NAT2* expression can therefore affect therapeutic response to drugs. Individuals with reduced *NAT2* expression may have different drug levels than those with normal expression. Consequently, they may need a different dosage schedule of drugs that are metabolized by *NAT2*.

INH acetylator status can be determined using phenotypic or genotypic methods. Humans can be phenotypically divided into "slow acetylators" and "rapid acetylators" and genotypically into three genotypes: homozygous slow (SS), heterozygous fast (FS), and homozygous fast (FF). The SS genotypes are those who have several mutations in the *NAT2* gene (*NAT2*5A*, *NAT2*5B* and *NAT2*6A*) [78]. They are at greater risk of drug-induced adverse reactions due to reduced drug elimination. The FS genotypes are those with one active and one inactive *NAT2* allele (heterozygous). They might require lower than average drug dosages for favorable therapeutic response. The FF genotypes are those in which the *NAT2*4* allele codes for a completely active enzyme. The polymorphic metabolism of INH is genetically determined in a simple Mendelian fashion, and the frequencies of the genes controlling slow or rapid INH metabolism vary among racial populations. The acetylator gene frequency for slow allele ranges from 10% in the Mongoloid race like the Eskimos, Japanese, and Chinese to 90% in Middle Eastern races, 60% in the Negroid, Caucasian, and south Indian populations, and 72% in the United States. Figure 36.9 shows the variation in plasma INH concentrations in the three genotypes of *NAT2* gene [79].

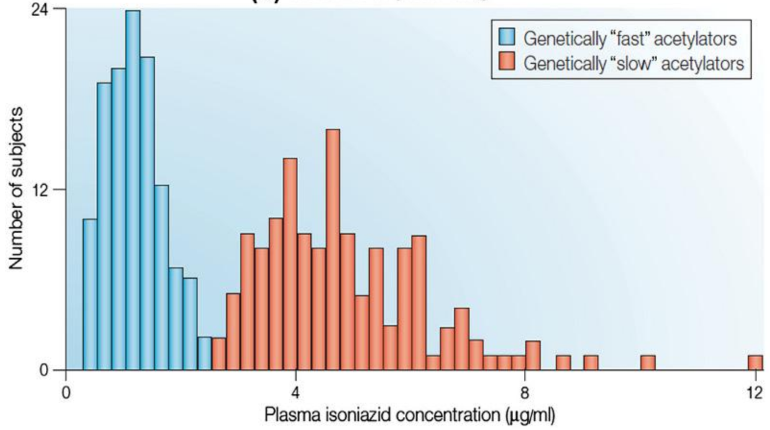
A study was undertaken by Kinzig-Schippers and others of 18 healthy Caucasians to assess individual INH exposure based on the *NAT2* genotype to predict personalized therapeutic dose [80]. They observed that *NAT2* genotypes accounted for 88% variability in apparent INH clearance, while INH preparation and body weight accounted for 2% and 3% of variations, respectively [80]. The researchers found a linear relationship between INH clearance and the number of high-activity *NAT2* alleles. The recommended standard daily dose of INH in adults is 5 mg/kg, but this can be roughly adjusted to 2.5 mg, 5.0 mg, and 7.5 mg/kg for subjects with, respectively, none (slow acetylators), one (intermediate acetylators), and two (rapid acetylators) alleles to achieve similar INH exposure. In another study, Parkin and colleagues reported differences in INH disposition among subjects with different *NAT2* genotypes, slow acetylators having four-fold to six-fold higher serum INH concentrations [81]. Based on this study, an individualized INH dosing regimen was suggested. Similarly, Kita and coworkers compared urinary INH excretion between healthy volunteers and patients, and found that subjects with a higher number of active *NAT2* alleles had higher levels of INH acetylation [82].

36.10.1.3 Isoniazid Acetylator Status and Treatment Efficacy

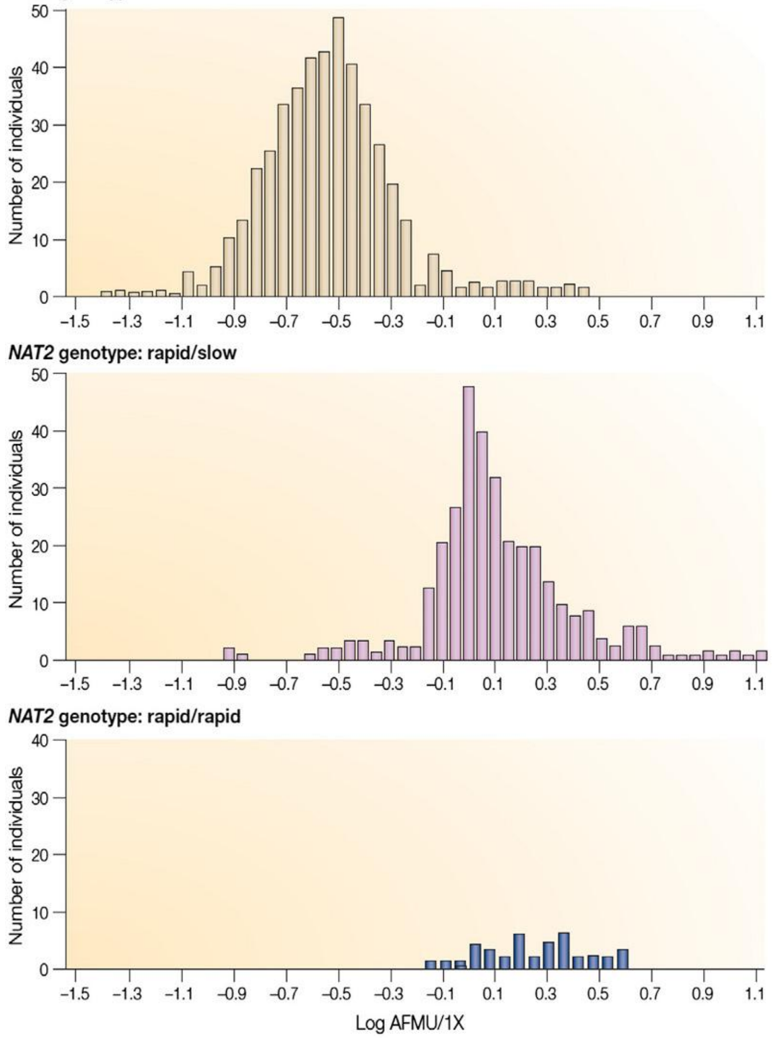
Limited information is available on the association between INH blood concentrations and efficacy and toxicity. Donald and others undertook a study to examine the influence of the human *NAT2* genotype on early INH bactericidal activity in 97 patients with sputum smear–positive pulmonary TB. Significant differences were found in the mean EBA activity of INH between slow and rapid acetylators and slow and intermediate acetylators. In this way, it was confirmed that *NAT2* genotypes exert a significant effect on INH EBA over a range of doses [83]. Thus, possible that low plasma INH levels in rapid acetylators are one of the causes of occasional therapeutic failure or relapse [84]. In contrast, higher levels in slow acetylators may predispose to toxicity. For these reasons, *NAT2* genotyping prior to INH administration may be clinically useful for predicting pharmacokinetic variability and for possible adjustment of INH dosing regimen.

In the 1950s and 1960s, Studies to understand the relationship between INH acetylator status and treatment outcome in patients with pulmonary TB receiving treatment with INH alone or in combination with p-amino salicylic acid were undertaken at the Tuberculosis Chemotherapy Centre in Chennai, India. Comparison of response to TB treatment between slow and rapid INH inactivators suggested an association between treatment response and rate of INH inactivation; there was a difference in the rate of conversion to bacteriological negativity between slow and rapid inactivators [85].

Under trial conditions, INH acetylator status had little prognostic significance during daily, twice-weekly, or thrice-weekly TB treatment [75]. However, under field conditions or in cases of drug irregularity, INH acetylator status may assume importance and may have an appreciable effect on outcomes among patients on twice- and thrice-weekly regimens [86]. Controlled clinical trials using once-weekly INH-containing regimens found that response to treatment was better in slow acetylators compared to rapid acetylators—the differences in cure rate were in the order of 20–35% ($p < 0.001$) [87]. To understand this inferior response of rapid acetylators, studies of serial serum INH concentrations in slow and rapid acetylators were undertaken. The peak concentration attained in slow acetylators was slightly higher than that in rapid acetylators; however, the fall-off was fast in rapid acetylators but more gradual in slow acetylators. Coverage, defined as the number of hours a bacteriostatic concentration of INH (0.2 μg/ml) is maintained in blood, was significantly greater in slow acetylators. Exposure, defined as the area under the time-concentration curve (AUC) was also significantly greater. Significant differences in coverage and exposure between slow and rapid acetylators suggest that, among rapid acetylators, failure of once-weekly INH regimens results from suboptimal coverage and exposure [88].



(a) Isoniazid as probe drug



(b) Caffeine as probe drug

FIGURE 36.9 **PK variation due to genetic variation in acetylation status.** Classic pharmacogenetic traits: inherited variation in N-acetylation. (a) Plasma concentrations of the antituberculosis agent isoniazid in 267 subjects 6 hours after an oral dose. The bimodal distribution results from polymorphisms in the gene encoding N-acetyltransferase-2 (NAT2), which catalyzes the metabolism of isoniazid. (b) NAT2 acetylation, measured as a ratio of the caffeine metabolites 5-acetylamino-6-formylamino-3-methyluracil (AFMU) and 1-methylxanthine (1X), in 795 unrelated German subjects. The antimode log AFMU/1X was −0.3 (10 – 0.3 = 0.5). 45.3% of the subjects were phenotypically rapid acetylators and 54.7% were phenotypically slow acetylators. NAT2 genotypes showed that 444 (55.8%) were slow/slow, 312 (39.2%) were rapid/slow, and 39 (4.9%) were rapid/rapid. Therefore, 5.7% of subjects were genotype–phenotype discordant. *Source: Reproduced with permission from Weinshilboum and Wang* [79].

A study by Weiner and others reported on the disadvantages that rapid acetylators have because of the development of rifamycin monoresistance in HIV-infected individuals receiving once-weekly intermittent therapy. This study showed poor treatment outcomes to be associated with INH acetylator status in patients on a once-weekly regimen [84].

The INH acetylator phenomenon may have important clinical implications in certain settings, such as during TB treatment in HIV-infected individuals, and in patients with chronic renal failure. A study undertaken by our group to understand the influence of HIV infection on the pharmacokinetics of anti-TB drugs showed that differences in INH peak concentration and exposure between HIV-infected and HIV-uninfected TB patients were more significant in rapid than in slow acetylators. Thus, among HIV-infected TB patients, where bioavailability of anti-TB drugs has been shown to be suboptimal, rapid acetylators are at a further disadvantage [89].

Determination of INH acetylator status may be particularly important in TB patients with chronic renal failure. The pharmacokinetics of INH was compared between patients with chronic renal failure and healthy volunteers. While plasma INH concentrations were similar in both groups, peak INH concentrations, exposure, and half-life were significantly higher in slow acetylators with renal failure than in healthy subjects ($p < 0.01$). This study suggested that it is necessary to reduce the dose of INH in slow acetylators to minimize adverse events [90].

In a study that examined the impact of INH dose and *NAT2* genotype and phenotype on the pharmacokinetics and pharmacodynamics of INH in adult pulmonary TB patients, it was observed that fast INH acetylators receiving a daily 6 mg/kg dose had exposure similar to that of slow acetylators who received a 3 mg/kg dose daily. It was suggested that reducing the dose below 6 mg/kg would be disadvantageous for fast acetylators; however, 3 mg/kg was sufficient for slow acetylators to achieve optimal INH exposure [91].

Wilkins and others designed a study to characterize the population pharmacokinetics of INH in South African pulmonary TB patients. A model based on data from 235 patients receiving oral doses of INH was developed to understand the pharmacokinetic variability. This study observed that the existing treatment guidelines might cause suboptimal exposure in rapid INH acetylators [92].

36.10.1.4 Isoniazid Acetylator Status in Pediatric TB

The pharmacokinetic study of Jeena and coworkers of children with TB suggested that INH doses may have to be individualized based on age, acetylator status, and disease process [93]. Other studies have also shown that decreased INH exposure was likely to impact TB treatment outcomes in children who were rapid acetylators compared to those who were slow acetylators [94–96]. Schaaf and others suggested that, in view of the enhanced rate of INH metabolism in young children, those below five years of age receive at least a 10 mg/kg dose to ensure that faster acetylators are exposed to adequate serum INH concentrations [95]. We undertook a pharmacokinetic study of anti-TB drugs in 84 children with TB. It was observed that INH acetylator status (AOR 4.2; 95% CI 1.1–15.4; $p = 0.033$) significantly impacted TB treatment outcome [96].

36.10.1.5 Isoniazid-Induced Hepatotoxicity

Isoniazid is quite nontoxic in conventional doses, with hepatotoxicity the most common and serious adverse reaction to occur during treatment. The incidence of hepatotoxicity during anti-TB treatment is reported to vary 2–28% [97]. This toxicity has been related to the slow clearance of acetyl hydrazine, a toxic metabolite of INH in affected patients who were identified as slow acetylators [68].

Hepatotoxicity is a major toxic effect of INH. This has caused concern, because INH is widely used in preventive therapy programs, apart from its use in TB treatment.

36.10.1.6 Mechanism of INH-Induced hepatotoxicity

INH-induced hepatotoxicity is considered to be idiosyncratic [98]. It is not the result of hypersensitivity or an allergic reaction [75,99] and is most probably caused by toxic metabolites. Although the predominant pathway of INH metabolism is acetylation by the hepatic enzyme NAT2, a small part of INH is directly hydrolyzed into isonicotinic acid and hydrazine; this pathway is of greater quantitative significance in slow than in rapid acetylators [100].

Most previous research focused on the hypothesis that acetyl hydrazine is the toxic metabolite of INH [99]. More recent studies, however, suggest that hydrazine, not INH or acetyl INH, is most likely to be the cause of INH-induced hepatotoxicity [100,101]. In early studies, it was suggested that fast acetylators are more susceptible to developing liver injury [99,102]. Again, however, more recent studies show that slow acetylators develop drug-induced liver injury more often and more severely when compared with fast acetylators [103,104]. In slow acetylators, more INH is left for direct hydrolysis into hydrazine; also, the "accumulated" acetyl hydrazine can be converted into hydrazine.

Mitchell and colleagues proposed that the incidence of hepatotoxicity due to INH may be greater in rapid than in slow acetylators because rapid acetylators produce monoacetyl hydrazine from INH faster than do slow acetylators; monoacetyl hydrazine is subsequently converted by the cytochrome P450-dependent hepatic microsomal enzyme system into a potent acylating agent that may cause hepatic necrosis [102]. However, this hypothesis was questioned because of the fact that formation of monoacetyl hydrazine was similar in the two phenotypes, its more rapid formation being offset by its more rapid elimination via diacetyl hydrazine in rapid acetylators [105].

RMP is a potent inducer of the hepatic microsomal enzyme system in the liver and intestine, thereby increasing the metabolism of many other compounds [106]. The combined use of INH and RMP has been associated with an increased risk of hepatotoxicity. Raghupati Sarma and others investigated the effect of daily RMP administration on the direct conversion of INH to isonicotinic acid and hydrazine by INH hydrolase in six slow and eight rapid acetylators [107]. They observed that the proportion of INH metabolized in slow acetylators through the direct pathway was 3% when INH alone was given compared to 6% during concomitant RMP administration. In rapid acetylators, the proportions were considerably less (<1% and 2.5%, respectively), suggesting that INH hydrolase was induced by RMP. In the direct pathway, INH is converted to isonicotinic acid and hydrazine. Thus, hydrazine, a known hepatotoxic agent in animals formed from INH, [108], was responsible for INH hepatotoxic action. The direct pathway of INH metabolism induced by RMP was shown to be more pronounced in slow than in rapid acetylators.

Most anti-TB drugs are liposoluble, and their elimination requires biotransformation into more water-soluble compounds. This is mostly performed by phase I and phase II biotransformation enzymes. In the phase I reaction, oxidation or demethylation occurs, usually performed by cytochrome P450 (CYP450) enzymes. The compound is still not very water soluble and requires further metabolism. Phase I reactions often produce toxic intermediates. In a typical phase II reaction, a large water-soluble compound is attached by glucuronidation or sulphation, resulting in nontoxic metabolites that can be easily eliminated. A third metabolic step for detoxification involves glutathione, which can covalently bind to toxic compounds by the enzyme glutathione S-transferase [109].

Human genetic studies have shown that cytochrome P450 2E1 (CYP2E1) is involved in anti-TB drug–induced liver toxicity [110,111]. The CYP2E1 cl/cl genotype is associated with a higher CYP2E1 activity and may lead to a higher production of hepatotoxins. Rat studies showed that isoniazid and hydrazine induce CYP2E1 activity [112,113]. INH has an inhibiting effect on CYP1A2, 2A6, 2C19, and 3A4 activity [114]. CYP1A2 is suggested to be involved in hydrazine detoxification [112,113]. INH can induce its own toxicity, possibly by the induction or inhibition of these enzymes. Thus, INH-induced hepatotoxicity is the interplay of acetylation by NAT2, oxidation by cytochrome P450 oxidase (CYP2EI), and detoxification by glutathione S-transferase (GST).

36.10.1.7 Genetic Susceptibility to Drug-Induced Liver Injury

Kim and others adopted a candidate gene approach to identify genetic susceptibility to anti-TB drug–induced hepatitis, by studying the association of genetic polymorphisms in phase I and II drug-metabolizing enzymes with anti-TB drug–induced hepatitis [115]. In 67 cases and 159 controls, the researchers compared genotype distributions of tagging SNPs in promoters, exons, and haplotypes in seven drug-metabolizing enzyme genes (*CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP2E1*, *NAT2*, *UGT1A1*, and *UGT1A3*). They observed that among four tagging SNPs of *NAT2*, −9796T > A in promoter and R197Q, were significantly associated ($p = 0.0016$ and $p = 0.0007$, respectively). *NAT2* haplotype 2 (A-A-A-G), carrying the A allele of −9796T > A and the A allele of R197Q, showed significant association ($p = 0.0004$). However, there was no significant association between genotypes of other enzyme-metabolizing genes and anti-TB drug–induced hepatitis. The constructs containing −9796A of *NAT2* showed significant lower luciferase activity ($p < 0.01$), suggesting decreased *NAT2* expression. The variant alleles and haplotype 2 showed significantly higher peak serum INH levels, a lower acetyl INH:INH ratio, and lower INH clearance, compared with wild types.

Kim and colleagues suggested that genetic variants in the promoter and exons of *NAT2* increase the risk of anti-TB drug–induced hepatitis by modifying acetylation phenotypes and/or gene *NAT2*expression, and that there is no essential role for genetic mutation of the other metabolizing enzymes in the development of this adverse reaction [115]. Further, their study showed that there were no significant differences in the levels of RMP, PZA, and EMB according to *NAT2* genotype.

Yimer and others [116] reported on the susceptibility of coinfected Ethiopian patients to drug-induced liver injury (DILI) when receiving efavirenz-based antiretroviral therapy (ART) for HIV in the face of RMP-based anti-TB therapy. The overall incidence of liver injury in this population was 30%, with more than half being severe cases. Yimer and colleagues' findings in a prospective case-control study indicated that patients with higher efavirenz plasma levels were at higher risk for DILI, with genetic predictors that included having the slow acetylator *NAT2* genotype and the *ABCB1* 3435TT genotype. Hence, it was suggested that pharmacogenetic testing be employed before initiation of ART and anti-TB therapies in order to identify the patients at risk.

36.10.1.8 NAT2 Polymorphisms

One study reported a 28-fold higher risk of INH-induced hepatotoxicity in slow compared to rapid acetylators in Japanese and Taiwanese populations [104]. Another study, of a Korean population, observed that *NAT2* slow acetylators had a 3–8-fold increase in this risk ($p = 0.005$). The latter study suggested that *NAT2* genotyping might serve as a useful tool to predict anti-TB drug–induced hepatotoxicity [117]. Slow *NAT2* acetylator status as a risk factor was observed in a study conducted in Tunisian patients with TB; it was concluded that diplotypes, *NAT2*53/5B*, *NAT2*6A/6A*, 481T/T, and 590A/A may serve as useful biomarkers predicting this condition [118]. A retrospective study of 102 Japanese TB patients showed that the occurrence of side effects was greater in slow acetylators than in rapid/intermediate acetylators [119]. A recent meta-analysis by Wang and coworkers of 14 studies, comprising 474 cases and 1446 controls, showed a significant association between *NAT2* slow acetylators and anti-TB-drug–induced liver toxicity risk [70]. The OR for *NAT2* slow acetylators compared with rapid acetylators was 4.7 (95% CI 3.3–6.7; $p < 0.001$). Subgroup analysis indicated that both Asian and non-Asian slow acetylators develop liver damage more frequently. On comparing *NAT2* intermediate acetylators and rapid acetylators, the OR was 1.3 (95% CI 0.93–1.71; $p < 0.14$). Another subgroup analysis for different treatment regimens showed that the combined ORs for INH + RMP, INH + RMP + PZA + EMB and INH were, respectively, 34.3 (95% CI 10.4–113.0; $p < 0.001$), 4.1 (95% CI 2.8–6.0; $p < 0.001$), and 2.4 (95% CI 0.52–10.73; $p < 0.27$) on comparing *NAT2* slow and rapid acetylators [70].

Contrary to these findings, a case-control study undertaken in TB patients of Caucasian origin did not find any association between increased risk of anti-TB drug–induced liver injury and the presence of slow *NAT2* polymorphisms [120]. Moreover, genotyping studies in an Indian population [121] and in a heterogenous population consisting of Hispanics, Africans, Caucasians, South Americans, and Asians [111] reported a lack of association between INH-induced hepatotoxicity and *NAT2* acetylator status.

36.10.1.9 CYP2E1 Polymorphisms

CYP2E1 is an enzyme of the hepatic microsomal enzyme system. It brings about conversion of acetyl hydrazine to hepatotoxins—namely, acetyl diazene, ketene, acetylonium ion, and the like [122]. It has been proposed that INH or its hydrazine metabolite may induce *CYP2E1* activity in rat; this activity is driven by blood INH levels [122]. On the other hand, INH may also inhibit *CYP2E1* activity, this being more prominent in individuals with the *CYP2E1* variant genotype [110]. Two point mutations in *CYP2E1*'s 5′ flanking region (PstI, RsaI sites) were found to be associated with higher transcription and increased enzyme activity [123]. Individuals with *c1* allele exhibited higher transcriptional activity, protein level, and enzymatic activity than those with the *c2* allele.

The DraI polymorphism has been reported to be located in intron 6 of *CYP2E1*, which has been shown to be linked with other polymorphic sites in the exonic region [77]. Further, the *c1/c1* genotype significantly increased the risk of anti-TB drug–induced hepatotoxicity when combined with the slow acetylator *NAT2* genotype. This was attributed to the fact that, in slow acetylators, less acetyl hydrazine is detoxified by NAT2-catalyzed acetylation and thus more acetyl hydrazine is oxidized by *CYP2E1* into toxic intermediates [68].

For a Chinese population, Lee et al. [124] reviewed 140 cases of patients treated for TB and found that slow acetylators defined by *NAT2* genotypes had a higher risk of anti-TB–induced liver injury as opposed to fast acetylators. They also concluded that *CYPE1* was associated with the severity of this condition. A study from Taiwan showed that the common **1A/*IA* genotype of *CYP2E1* enhanced the odds of INH-induced liver toxicity by 2.5 times in adult TB patients [110]. Further, a combination of the **1A/*1A* genotype and *NAT2* slow acetylation status increased these odds (OR 7.43; 95% CI 2.4–22.8) compared with the presence of any one polymorphism. Another study in Indian pediatric patients observed that the risk of liver toxicity was enhanced by the variant *CYP2E1*6* allele and the **IA*6*ID* haplotype at *CYP2E1* (OR 11.0; 95% CI 1.02–11.0 and OR: 4.6; 95% CI 1.3–16.3, respectively) [125]. The common **IA* allele at *CYP2E1* that increases liver enzyme levels was observed in a heterogenous group of Asians, Africans, Caucasians, South Americans, Hispanics, and Africans [120]. However, a study of a Korean population did not find any such association [117].

Teixeira and others [126] studied *NAT2*, *CYP2E1*, and *GST* gene polymorphisms in 167 patients undergoing TB treatment. The slow acetylators had a higher incidence of hepatitis than the moderate or rapid acetylators, as related to polymorphisms in those specific genetic loci. An and others [127] also looked at *NAT2*, as well as the P450 *CYP2E1* mentioned previously. Their results also revealed that subjects with *NAT2* alleles had an increased susceptibility to liver injury from anti-TB medication—specifically, those who had the slow acetylator type *NAT2*. They also found the *CYP2E1* genotype to be significantly involved with susceptibility to hepatotoxicity unless it was combined with the slow *NAT2* acetylator.

Another study by An and others [128] evaluated 101 patients who experienced anti-TB drug–induced liver injury and revealed that those with the 282 T/T, 590A/A, 857 G/A, or *857A/A* genotypes were more likely to develop it. Variant diplotypes *NAT2*5B/5B*, and *NAT2-6A/6A* were found by Ben Mahmoud and coworkers [118], in their study of 14 patients with anti-TB drug–induced hepatotoxicity, to be more frequent and that, overall, slow acetylator *NAT2* alleles were more susceptible to hepatotoxicity from these drugs.

Not all investigators have found such consistent observations. For example, Tang et al. [129] evaluated 4304 patients who had undergone anti-TB therapy and found no statistically significant positive association with the *CYP2E1* or *GSTM1* and *GSTT1* genetic polymorphisms.

36.10.1.10 GSTM1, GSTT1 Polymorphisms

The GSTs are a group of enzymes that are involved in the detoxification of carcinogens, toxic chemicals, and medications [130]. They bring about the conjugation of glutathione and the substrate for solubilization and elimination from the body. The GST enzymes are coded by at least five distinct loci. Of these, *GSTM1* and *GSTT1* are reported to be associated with hepatotoxicity [131]. GSTs are reported to play a key role in INH metabolism. Glutathione serves as an intracellular scavenger of free radicals by conjugating with toxic INH metabolites. Inadequate GST activity due to homozygous deletion at the *GSTM1* and *GSTT1* loci can influence susceptibility to INH-induced liver toxicity.

The risk of INH-induced hepatotoxicity has been shown to be increased (OR 2.13; 95%CI 1.25-3.10) in Indian patients with the *GSTM1* homozygous mutation [121]. Also, while Taiwanese patients with *GSTM1* homozygous deletion had twice the risk of this condition, such risk did not occur in those with the *GSTT1* homozygous deletion [117,131]. Similar studies with reasonable sample sizes should be carried out in different ethnic populations. Identifying patients with the *GSTM1* homozygous deletion may help in preventing DILI and better management of it when it occurs.

36.10.1.11 UDP Polymorphisms

Chang and coworkers [132] examined the association of genetic variations of UDP glucuronosyl transferase (A) (UGT1A1) and anti-TB drug–induced hepatotoxicity. In their series of 17 patients, UGT1A1*28 compounds were related to increased DILI risk.

36.10.1.12 TNF-Alpha Polymorphisms

Tumor necrosis factor (TNF) -alpha plays an important role in immune responses [133]. Kim et al. [134] studied its impact on hepatotoxicity. Seventy seven patients with anti-TB drug–induced live injury in their study were found to have a higher frequency of the genotype TNF-alpha-308 G/A compared to patients on the same drug regimen who did not develop hepatotoxicity.

36.10.1.13 Manganese Superoxide Dismutase

It is known that hepatic damage can be caused by reactive oxygen species (ROS) that are formed from various reactions. Manganese superoxide dismutase (MnSOD) brings about a decrease in ROS load in mitochondria. A study from Taiwan showed that genotypes with the variant C allele (the T > C polymorphism at codon 47; alanine > valine) enhanced the risk of anti-TB drug–induced liver toxicity [131]. This is probably because, along with the variant amino acid, valine, at codon 47, it augmented the formation of toxic hydrogen peroxide, which can lead to DILI.

36.10.1.14 Human Leucocyte Antigen Alleles

The lack of HLADQA1*0102 and the existence of the DQB1*0201 alleles have been reported to have an independent association with enhanced risk of anti-TB–induced hepatotoxicity in Indian patients [135].

36.10.2 Rifampicin

Rifampicin (RMP) is an important first-line drug used to treat TB. It has demonstrated concentration-dependent activity against *M. tuberculosis* under *in vitro* and *in vivo* conditions [136]. Marked intersubject variations in RMP levels are reported to occur, but the basis for variation has not been well characterized. Weiner and colleagues undertook a pharmacokinetic study to explore why huge inter-individual variations in RMP levels are seen in patients receiving standard treatment dosages [72].

Organic anion transporter peptides (OATPs) play an important role in the transport and disposition of drugs in the human system. RMP is a substrate of these drug transporters, which include p-glycoprotein and OATP1B1. Nuclear receptors—namely, pregnane X receptor (PXR) and constitutive androstane receptor (CAR)—transcriptionally regulate drug transporters. The pharmacokinetics of RMP was compared between TB patients and healthy subjects, between regions (African and non-African), and between races (black and all others).

The relationship between RMP pharmacokinetics and polymorphisms of drug influx and efflux transporter genes—anion-transporting polypeptides (*SLCO1B1*-463C > A and 521T > C and *SLCO1B3*) and p-glycoprotein (*ABCB1*)—was also studied [72]. The findings from a study of 72 adult pulmonary TB patients from North America, Africa, and Spain showed that polymorphisms in the *SLCO1B1* gene had a significant influence on RMP exposure: 36% lower among SLCO1B1 463 CA genotypes than among CC genotypes (29.8 versus 46.7 μg.h/ml; $p = 0.001$) [72]. Further, *SLCO1B1* gene polymorphisms related to reduced RMP exposure were more pronounced in black subjects. This study was the first to report an association between RMP exposure and *SLCO1B1* 463 C>A gene polymorphism. Marked interindividual RMP exposure variations were probably due to this polymorphism.

Chigutsa and others undertook a study of South African TB patients to examine the impact of *ABCB1*, *SLCO1B1*, *PXR*, and *CAR* polymorphisms on RMP pharmacokinetics [73]. Patients heterozygous and homozygous for the variant allele of *SLCO1B1* rs4149032 polymorphism, showed decreases in RMP bioavailability of 20% and 28%, respectively. This polymorphism explained 21% of the between-subject variability in drug clearances. Simulations revealed that an increased RMP dose of approximately 30% in patients harboring the polymorphism resulted in plasma RMP levels similar to those with the wild-type allele, with a typical peak concentration above the recommended minimum of 8 μg/ml. The other polymorphisms—*ABCB1*, *PXR*, and *CAR*—however, did not have any significant impact on RMP pharmacokinetics. This study suggested that an increase in RMP dose is desirable for *SLCOB1B1* carriers. More studies in different populations are required to understand the role of *SLCOB1B1* polymorphisms in influencing RMP plasma concentrations and their role in influencing TB treatment outcomes.

36.10.3 Aminoglycosides

Aminoglycosides (e.g., streptomycin, kanamycin, and amikacin) are used in TB treatment. Ototoxicity is an adverse event that may result from the use of aminoglycosides. Mutations in mitochondrial DNA (mt DNA) have been found to be associated with both aminoglycoside-induced and nonsyndromic deafness. The A1555G mutation has also been found in a number of families and patients with sporadic aminoglycoside-induced severe or profound hearing loss [74,137]. It lies in the conserved decoding region of small ribosomal RNA, which is important for aminoglycosides action [138].

In fact, the new C1494-1555G pair in 12S rRNA created by the A1555G mutation facilitates the binding of aminoglycosides to mitochondrial 12S rRNA [139]. Functional studies demonstrated a decrease in the growth rate of lymphoblastoid cells derived from symptomatic and asymptomatic members of the Arab-Israel pedigree in the presence of a high concentration of neomycin or paromomycin [140]. In addition, a C insertion or deletion at position 961 of the 12S rRNA gene has been shown to be associated only with aminoglycoside-induced deafness [141]. These data strongly indicate that the human mitochondrial 12S rRNA, particularly that carrying the A1555G mutation, is the main ototoxic target for aminoglycoside antibiotics. Mutations at positions 1555 and 961 in the 12S rRNA gene account for only a small portion of patients with aminoglycoside-induced hearing impairment.

Zhao and others undertook a molecular characterization study in a Chinese family with aminoglycoside-induced and nonsyndromic hearing loss to understand the association between aminoglycoside-induced ototoxicity and mitochondrial mutations. It was observed that a C>T transition at position 1494 in the mitochondrial 12S rRNA gene was responsible in individuals genetically predisposed to ototoxicity [74]. This mutation is localized at the A site of the mitochondrial 12S rRNA [138,142], which has been implicated to be associated with aminoglycoside ototoxicity [139]. It is likely that the deafness phenotype linked to the C>T 1494 12S rRNA gene polymorphism can be induced or made worse by aminoglycosides.

36.10.4 Predicting Drug–Induced Liver Injury

Considerable research continues to be directed toward identifying genetic and other risk factors for drugs causing idiosyncratic drug-induced liver injury [143]. Among the most commonly reported drugs causing severe drug-induced liver are microbial agents such as INH and other anti-TB drugs. Being able to predict individuals who break down with injury remains a major research focus. Classic clinical predictors were studied by Singla and others [144], who evaluated 175 cases of anti-TB drug–induced liver injury and found that subjects above 35 years of age and with poor nutritional status were independent risk factors.

36.11 CONCLUSION

Although genetic polymorphisms are known to impact drug effects, individualized drug therapy is seldom used in clinical practice. The cornerstone of TB management is a six-month course of INH, RMP, PZA, and EMB. Compliance is crucial for curing TB. Adverse events often negatively affect compliance, and this can impact treatment outcome. As for anti-TB drugs, genetic factors are known to influence INH metabolism and disposition; this alters INH levels in different patients to varying degrees and puts some of them at greater risk for liver toxicity. Determining INH acetylator status will aid clinicians in identifying patients at higher DILI risk. Close monitoring of these patients might prove useful in reducing the occurrence of adverse drug effects and enhancing treatment success. In the long run, it will probably decrease healthcare costs.

Pharmacogenomic studies have the potential to achieve this by translating genome variability into improved therapeutics. Importantly large clinical trials will be needed to prospectively evaluate the merits of genotyping and of adjusting TB drug dose during treatment. These studies must link pharmacogenetic data with TB treatment outcomes.

The US Food and Drug Administration now requires genetic testing prior to treatment with abacavir and carbamazepine (in patients of Han Chinese or Thai origin) and with a few other drugs for cancer therapy [145]. This shows that pharmacogenetics is becoming more common in clinical practice. New developments in genotyping technologies, including next-generation sequencing and epigenetics will move the field of pharmacogenetics forward. However, one has to keep in mind the cost-effectiveness and feasibility of integrating pharmacogenomics in routine practice. It is important to establish the benefits of genetic testing and individualized treatment before it becomes a routine exercise.

REFERENCES

[1] Dye C, Watt CJ, Bleed DM, Hosseini SM, Raviglione MC. Evolution of tuberculosis control and prospects for reducing tuberculosis incidence, prevalence, and deaths globally. JAMA 2005;293:2767–75.

[2] World Health Organization. Global TB report; 2012.

[3] Lawn SD, Zumla AI. Tuberculosis. Lancet 2011;378(9785):57–72.

[4] Tiemersma EW, Van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJ. Natural history of tuberculosis: duration and fatality

of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. PLoS One 2011;6:e17601.

[5] Sharma SK, Mohan A. Extrapulmonary tuberculosis. Indian J Med Res 2004;120:316–53.

[6] Marais BJ. Childhood tuberculosis: epidemiology and natural history of disease. Indian J Pediatr 2011;78:321–7.

[7] Raviglione M, Marais B, Floyd K, Lönnroth K, Getahun H, Migliori GB, et al. Scaling up interventions to achieve global TB control: progress and new developments. Lancet 2012;379(9829):1902–13.

[8] International Standards for Tuberculosis Care. <http://www.who.int/tb/publications/2006/istc_report.pdf> [accessed 6.3.13].

[9] Luelmo F. What is the role of sputum microscopy in patients attending health facilities? In: Frieden T, editor Toman's tuberculosis: case detection, treatment, and monitoring—questions and answers. 2nd ed.. Geneva: World Health Organization; 2004. p. 7–13.

[10] Rawat J, Biswas D, Sindhwani G, Kesharwani V, Masih V, Chauhan BS. Diagnostic defaulters: an overlooked aspect in the Indian revised national tuberculosis control program. J Infect Dev Countries 2012;6:20–2.

[11] Rawat J, Biswas D, Sindhwani G, Kesharwani V, Masih V. An alternative 1-day smear microscopy protocol for the diagnosis of pulmonary tuberculosis. Respirology 2010;15:1127–30.

[12] Cattamanchi A, Davis JL, Worodria W, den Boon S, Yoo S, Matovu J, et al. Sensitivity and specificity of fluorescence microscopy for diagnosing pulmonary tuberculosis in a high HIV prevalence setting. Int J Tuberc Lung Dis 2009;13:1130–6.

[13] Prabha Desikan. Sputum smear microscopy in tuberculosis: is it still relevant? Indian J Med Res 2013;137:442–4.

[14] Lawn SD, Mwaba P, Bates M, Piatek A, Alexander H, Marais BJ, et al. Advances in tuberculosis diagnostics: Xpert® MTB/RIF. Assay and future prospects for a point-of-care test. Lancet Infect Dis 2013;13(4):349–61. http://dx.doi.org/10.1016/S1473-3099(13)70008-2.

[15] Sharma SK, Mohanan S, Sharma A. Relevance of latent TB infection in areas of high TB prevalence. Chest 2012;142:761–73.

[16] Pinto LM, Grenier J, Schumacher SG, Denkinger CM, Steignart KR, Pai M. Immunodiagnosis of tuberculosis: state of the art. Med Princ Pract 2012;21:4–13.

[17] Rangaka MX, Willinson KA, Glynn JR, Ling D, Menzies D, Mwansa-Kambafwile J, et al. Predictive value of interferon-γ release assays for incident active tuberculosis: a systematic review and meta-analysis. Lancet Infect Dis 2012;12:45–55.

[18] World Health Organization. Use of tuberculosis interferon-gamma release assays (IGRAs) in low and middle income countries: policy statement WHO/HTM/TB/2011.18. Geneva: World Health Organization; 2011.<http://whqlibdoc.who.int/publications/2011/9789241502672_eng.pdf> [accessed 28.2.13]

[19] Gopi A, Madhavan SM, Sharma SK, Sahn SA. Diagnosis and treatment of tuberculous pleural effusion in 2006. Chest 2007;131:880–9.

[20] Xu HB, Jiang RH, Li L, Sha W, Xiao HP. Diagnostic value of adenosine deaminase in cerebrospinal fluid for tuberculous meningitis: a meta-analysis. Int J Tuberc Lung Dis 2010;14:1382–7.

[21] World Health Organization. Commercial serodiagnostic tests for diagnosis of tuberculosis: policy statement WHO/HTM/TB/2011.5. Geneva: World Health Organization; 2011.

[22] Jeong YJ, Lee KS. Pulmonary tuberculosis: up-to-data imaging and management. Am J Roentgenol 2008;191:834–44.

[23] Fox W, Mitchison DA. Short course chemotherapy for pulmonary tuberculosis. Am Rev Respir Dis 1975;111:845–8.

[24] Ma Z, Leinhardt C, Mcllleron H, Nunn AJ, Wang X. Global TB drug development pipeline: the need and the reality. Lancet 2010;375(9731):2100–9.

[25] Zumla A, Nahid P, Cole ST. Advances in the development of new TB drugs and treatment regimens. Nat Rev Drug Discov 2013;12:388–404.

[26] Mitchison D. Basic mechanisms of chemotherapy. Chest 1979;76(Suppl 6):771–81.

[27] David HL. Drug resistance in Mycobacterium tuberculosis and other mycobacteria. Clin Chest Med 1980;1:227–30.

[28] Iseman MD, Madsen LA. Drug resistant tuberculosis. Clin Chest Med 1989;10:341–53.

[29] Singapore Tuberculosis Service, British Medical Research Council. Five years follow up of a clinical trial of three 6-month regimens of chemotherapy given intermittently in the continuation phase in the treatment of pulmonary tuberculosis. Am Rev Respir Dis 1988;137:1147–50.

[30] Snider Jr DE, Graczyk J, Bek E, Rogowski J. Supervised 6-month treatment of newly diagnosed pulmonary tuberculosis using isoniazid, rifampin, and pyrazinamide with and without streptomycin. Am Rev Respir Dis 1984;130:1091–4.

[31] Hong Kong Chest Service, British Medical Research Council Study. Controlled trial of 6-month and 8-month regimens in the treatment of pulmonary tuberculosis—the results up to 24 months. Tubercle 1979;60:201–10.

[32] Combs D, O'Brien R, Geiter L. USPHS tuberculosis short course chemotherapy trial 21: effectiveness, toxicity and acceptability—the report of final results. Ann Intern Med 1990;112:397–406.

[33] Ben Amor Y, Day MS, Schluger NW. Preventing the next generation of extensively drug-resistant tuberculosis. Int J Tuberc Lung Dis 2010;14:525–7.

[34] Gandhi NR, Nunn P, Dheda K, Schaaf HS, Zignol M, van Soolingen D, et al. Multidrug resistant and extensively drug resistant TB: a threat to global control of TB. Lancet 2010;375(9728):1830–43.

[35] Dye C, Williams BG, Espinal MA, Raviglione MC. Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. Science 2002;295:2042–6.

[36] Farmer P, Kim JY. Community based approaches to the control of multidrug resistant tuberculosis: introducing 'DOTS-Plus'. Br Med J 1998;317:671–4.

[37] Crofton J, Chaulet P, Maher D. Guidelines for the management of drug-resistant tuberculosis WHO/TB/96.210 (Rev 1). Geneva: WHO; 1997.

[38] Jain S, Rodrigues C, Mehta A, Udwadia ZF. High prevalence of XDR-TB from a tertiary care hospital in India. Am J Respir Crit Care Med 2007;175:A510.

[39] Datta BS, Hassan G, Kadri SM, Qureshi W, Kamili MA, Singh H, et al. Multidrug-resistant and extensively drug resistant tuberculosis in Kashmir, India. J Infect Dev Countries 2009;4:19–23.

[40] Murray M, Oxlade O, Lin HH. Modeling social, environmental and biological determinants of tuberculosis. Int J Tuberc Lung Dis 2011;15(Suppl 2):S64–70.

[41] Moller N, de Wit E, Hoal EG. Past, present and future directions in human genetic susceptibility to tuberculosis. FEMS Immunol Med Microbiol 2010;58:3–26.

[42] Najmi N, Kaur G, Sharma SK, Mehra NK. Human Toll-like receptor 4 polymorphisms TLR4 Asp299Gly and Thr399Ile influence susceptibility and severity of pulmonary tuberculosis in the Asian Indian population. Tissue Antigens 2010;76:102–9.

[43] Padmapriyadarsini C, Narendran G, Swaminathan S. Diagnosis & treatment of tuberculosis in HIV co-infected patients. Indian J Med Res 2011;134:850–65.

[44] Whalen C, Horsburgh CR, Hom D, Lahart C, Simberkoff M, Ellner J. Accelerated course of human immunodeficiency virus infection after tuberculosis. Am J Respir Crit Care Med 1995;151:129–35.

[45] Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC, et al. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch Intern Med 2003;163:1009–21.

[46] Narain JP, Raviglione MC, Kochi A. HIV associated tuberculosis in developing countries: epidemiology and strategies for prevention. Tuber Lung Dis 1992;73:311–21.

[47] Centers for Disease Control (CDC). Transmission of multidrug resistant tuberculosis among immunocompromized persons in a correctional system—New York, 1991. Morb Mortal Wkly Rep 1992;41:507–9.

[48] Beck-Sague C, Dooley SW, Hutton MD, Otte J, Breeden A, Crawford JT, et al. Hospital outbreak of multidrug resistant *Mycobacterium tuberculosis* infections. Factors in transmission to staff and HIV-infected patients. JAMA 1992;268:1280–6.

[49] Swaminathan S, Paramasivan CN, Ponnuraja C, Illiayas S, Rajasekaran S, Narayanan PR. Anti-tuberculosis drug resistance in patients with HIV and tuberculosis in south India. Int J Tuberc Lung Dis 2005;9:896–900.

[50] Jeon CY, Murray MB. Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. PLoS Med 2008;5:e152.

[51] Stevenson CR, Forouhi NG, Roglic G, Williams BG, Lauer JA, Dye C, et al. Diabetes and tuberculosis: the impact of the diabetes epidemic on tuberculosis incidence. BMC Public Health 2007;7:234.

[52] Pai M, Mohan A, Dheda K, Leung CC, Yew WW, Christopher DJ, et al. Lethal interaction: the colliding epidemics of tobacco and tuberculosis. Expert Rev Infect Ther 2007;5:385–91.

[53] Slama K, Chiang CY, Enarson DA, Hassmiller K, Fanning A, Gupta P, et al. Tobacco and tuberculosis: a qualitative systematic review and meta-analysis. Int J Tuberc Lung Dis 2007;11:1049–61.

[54] Styblo K, Sutherland I. Epidemiology of tuberculosis in children. Bull Int Union Tuberc 1982;57:133–9.

[55] Rouilon A, Perdrizet S, Parrot R. Transmission of tubercle bacilli. Tubercle 1976;57:275–99.

[56] Miller FJW, Seale RME, Taylor MD. In: Tuberculosis in children: evolution, control, treatment. London: Churchill; 1963. p. 73. [chapter 4]

[57] Lincoln E. Course and prognosis of tuberculosis in children. Am J Med 1950;9:623–32.

[58] Wallgren A. Primary pulmonary tuberculosis in childhood. Am J Dis Child 1935;49:1105–36.

[59] Starke JR, Jacobs RF, Jerb J. Resurgence of tuberculosis in children. J Pediatr 1992;120:839–55.

[60] World Health Organization. Treatment of tuberculosis. Guidelines for national programmes; 1993.

[61] Schaaf HS, Gie RP, Kennedy M, Beyers N, Hesseling PB, Donald PR, et al. Evaluation of young children in contact with adult multidrug-resistant pulmonary tuberculosis: a 30-month follow-up. Pediatrics 2002;109:765–71.

[62] Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. Science 1999;286:487–91.

[63] Evans WE, Mcleods HL. Pharmacogenomics—drug disposition, drug targets and side effects. N Engl J Med 2003;348:538–49.

[64] Vesell ES. Pharmacogenetic perspectives gained from twin and family studies. Pharmacol Ther 1989;41:535–52.

[65] Kalow W, Tang BK, Endreyi I. Hypothesis: comparisons of inter and intra individual variations can substitute for twin studies in drug research. Pharmacogenetics 1998;8:283–9.

[66] Mcleod HL, Evans WE. Pharmacogenomics: unlocking the human genome for better drug therapy. Annu Rev Pharmacol Toxicol 2001;41:101–21.

[67] Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001;409:928–33.

[68] Roy PD, Majumder M, Roy B. Pharmacogenomics of anti-TB drug resistant hepatotoxicity. Pharmacogenomics 2008;9(3):311–21.

[69] Tostmann A, Boeree MJ, Aarnoutse RE, de Lange WC, van der Ven AJ, Dekhuijzen R, et al. Antituberculosis drug-induced hepatotoxicity: concise up-to-date review. J Gasteroenterol Hepatol 2008;23(2):192–202.

[70] Wang PY, Xie SY, Hao Q, Xiang BF. NAT2 polymorphisms and susceptibility to anti tuberculosis drug-induced liver injury: meta-analysis. Int J Tuberc Lung Dis 2012;16(5):589–95.

[71] Sun F, Chen Y, Xiang Y, Zhan S. Drug-metabolizing enzyme polymorphism and predisposition to anti-tuberculosis drug-induced liver injury: a meta-analysis. Int J Tuberc Lung Dis 2008;12(9):994–1002.

[72] Weiner M, Peloquin CA, Burman W, Luo CC, Engle M, Prihoda TJ, et al. Effects of tuberculosis, race and human gene *SLCO1B1* polymorphisms on rifampin concentrations. Antimicrob Agents Chemother 2010;54(10):4192–200.

[73] Chigutsa E, Visser ME, Swart EC, Denti P, Pushpakom S, Egan D, et al. The SLCO1B1 rs419032 polymorphism is highly prevalent in South Africa and is associated with reduced rifampin concentration: dosing implications. Antimicrob Agents Chemother 2011;55(9):4122–7.

[74] Zhao H, Li R, Wang Q, Yan Q, Deng JH, Han D, et al. Maternally inherited aminoglycoside-induced and nonsyndromic deafness is associated with the novel C1494T mutation in the mitochondrial 12S rRNA gene in a large Chinese family. Am J Hum Genet 2004;74:139–52.

[75] Ellard GA, Gammon PT. Pharmacokinetics of isoniazid metabolism in man. J Pharmacokinet Biopharmaceut 1976;4:83–113.

[76] Peters JH, Miller KS, Brown P. Studies on the metabolic basis for the genetically determined capacities for isoniazid inactivation in man. J Pharmacol Exp Ther 1965;150:298–304.

[77] Metushi IG, Cai P, Zhu X, Nakagawa T, Utrecht JP. A fresh look at the mechanism of isoniazid-induced hepatotoxicity. Clin Pharmacol Ther 2011;89(6):911–4.

[78] Cascorbi L, Roots I. Pitfalls in N-acetyl transferase 2 genotyping. Pharmacogenetics 1999;9:123–7.

[79] Weinshilboum R, Wang L. Pharmacogenomics: bench to bedside. Discov Med 2005;5(25):30–6.

[80] Kinzig-Schippers M, Tomalik-Scharte D, Jetter A, Scheidel B, Jakob V, Rodamer M, et al. Should we use N-acetyl transferase type 2 genotyping to personalize isoniazid doses? Antimicrobial Agents Chemother 2005;49:1733–8.

[81] Parkin DP, Vandenplas S, Botha FJ, Vandenplas ML, Seifart HI, van Helden PD, et al. Trimodality of isoniazid elimination: phenotype and genotype in patients with tuberculosis. Am J Respir Crit Care Med 1997;155:1717–22.

[82] Kita T, Tanigawara Y, Chikazawa S, Hatanaka H, Sakaeda T, Komada F, et al. N-acetyl transferase-2 genotype correlated with

isoniazid acetylation in Japanese tuberculous patients. Biol Pharm Bull 2001;24:544–9.

[83] Donald PR, Sirgel FA, Venter A, Parkin DP, Seifart HI, Van de Wal BW, et al. The influence of human N-acetyl transferase genotype on the early bactericidal activity of isoniazid. Clin Infect Dis 2004;39:1425–30.

[84] Weiner M, Burman W, Vernon A, Benator D, Peloquin CA, Khan A, et al. Low isoniazid concentrations and outcome of tuberculosis treatment with once-weekly isoniazid and rifapentine. Am J Respir Crit Care Med 2003;167:1341–7.

[85] Selkon JB, Wallace Fox, Gangadharam PRJ, Ramachandran K, Ramakrishnan CV, Velu S. Rate of inactivation of isoniazid in south Indian patients with pulmonary tuberculosis-2. Clinical implications in the treatment of pulmonary tuberculosis with isoniazid either alone or in combination with PAS. Bull WHO 1961;25:779–92.

[86] Raghupati Sarma G. Isoniazid metabolism in man and its therapeutic implications. In: Proceedings of the symposium on clinical pharmacokinetics; 1978.

[87] . Controlled comparison of two fully supervised once weekly regimens in the treatment of newly diagnosed pulmonary tuberculosis. Tubercle 1973;54(1):23–45.

[88] Sarma GR, Kailasam S, Mitchison DA, Nair NG, Radhakrishna S, Tripathy SP, et al. Studies of serial plasma isoniazid concentrations with different doses of a slow release preparation of isoniazid. Tubercle 1975;56:314–23.

[89] Gurumurthy P, Ramachandran G, Hemanth Kumar AK, Rajasekaran S, Padmapriyadarsini C, Swaminathan S, et al. Decreased bioavailability of rifampin and other anti-TB drugs in patients with advanced HIV disease. Antimicrob Agents Chemother 2004;48:4473–5.

[90] Prema Gurumurthy S, Raghupati Sarnia G, Jayasankar K, Thyagarajan K, Prabhakar R, Muthusethupathi MA, et al. Single-dose pharmacokinetics of isoniazid and rifampicin in patients with chronic renal failure. Ind J Tub 1992;39:221–8.

[91] Donald PR, Parkin DP, Seifart HI, Schaaf HS, van Helden PD, Werely CJ, et al. The influence of dose and N-acetylransferase-2 (NAT2) genotype and phenotype on the pharmacokinetics and pharmacodynamics of isoniazid. Pharmacokinet Dispos 2007;63(7):633–9.

[92] Wilkins JJ, Langdon G, McIlleron H, Pillai G, Smith PJ, Simonsson US. Variability in the population pharmacokinetics of isoniazid in South African tuberculosis patients. Br J Clin Pharmacol 2011;72(1):51–62.

[93] Jeena PM, Bishai WR, Pasipanodya JG, Gumbo T. *In silico* children and the glass mouse model; clinical trial simulations to identify and individualize optimal isoniazid doses in children with tuberculosis. Antimicrobial Agents Chemother 2011;55:539–45.

[94] McIIleron H, Willemse M, Werely CJ, Hussey GD, Schaaf HS, Smith PJ, et al. Isoniazid plasma concentrations in a cohort of south African children with tuberculosis: implications for international pediatrics dosing guidelines. Clin Infect Dis 2009;48:1547–53.

[95] Schaaf HS, Parkin DP, Seifart HI, Werely C, Hesseling P, van Helden PD, et al. Isoniazid pharmacokinetics in children treated for respiratory tuberculosis. Arch Dis Child 2005;90:614–8.

[96] Ramachandran G, Hemanth Kumar AK, Bhavani PK, Poorana Gangadevi N, Sekar L, Vijayasekaran D, et al. Age, nutritional status and INH acetylator status affect pharmacokinetics of anti-tuberculosis drugs in children. Int J Tuberc Lung Dis 2013;17(6):800–6.

[97] Devadatta S, Gangadharam PRJ, Andrews RH, Fox W, Ramakrishnan CV, Selkon JB, et al. Peripheral neuritis due to isoniazid. Bull WHO 1960;23:587–98.

[98] Black M, Mitchell JR, Zimmerman HJ, Ishak KG, Epler GR. Isoniazid-associated hepatitis in 114 patients. Gastroenterology 1975;69:289–302.

[99] Mitchell JR, Zimmerman HJ, Ishak KG, Thorgeirsson UP, Timbrell JA, Snodgrass WR, et al. Isoniazid liver injury: clinical spectrum, pathology and probable pathogenesis. Ann Intern Med 1976;84:181–92.

[100] Sarich TC, Yousseti M, Zhou T, Adams SP, Wall RA, Wright JM. Role of hydrazine in the mechanism of isoniazid hepatotoxicity in rabbits. Arch Toxicol 1996;70:835–40.

[101] Noda A, Hsu KY, Noda H, Yammamoto Y, Kurozumi T. Is isoniazid-hepatotoxicity-induced by the metabolite, hydrazine? JUOEH 1983;5:183–90.

[102] Mitchell Jr Thorgeirsson UP, Black M, Timbrell JA, Snodgrass WR, Potter WZ, et al. Increased incidence of isoniazid hepatitis in rapid acetylators: possible relation to hydrazine metabolites. Clin Pharmacol Ther 1975;18:70–9.

[103] Pande JN, Singh SP, Khilnani GC, Khilinani S, Tandon RK. Risk factors for hepatotoxicity from antituberculosis drugs: a case control study. Thorax 1996;51:132–6.

[104] Ohno M, Yamaguchi I, Yamamoto I, Fukuda T, Yokota S, Maekura R, et al. Slow N-acetyl transferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity. Int J Tuberc Lung Dis 2000;4:256–61.

[105] Ellard GA. Slow-release preparation of isoniazid: pharmacological aspects. Bull Int Union Tuberc 1976;51:143–54.

[106] Kolars JC, Schmiedlin-Ren P, Schutz JD, Fang C, Watkins PB. Identification of rifampin-inducible P450IIIA4 (CYP3A4) in human small bowel enterocytes. J Clin Invest 1992;90:1871–8.

[107] Raghupati Sarma G, Immanuel Chandra, Kailasam S, Narayana ASL, Venkatesan P. Rifampin induced release of hydrazine from isoniazid: a possible cause of hepatitis during treatment of tuberculosis with regimens containing isoniazid and rifampin. Am Rev Respir Dis 1986;133:1072–5.

[108] Back KC, Thomas AA. Aerospace problem in pharmacology and toxicology. Annu Rev Pharmacol 1970;10:395–412.

[109] Lee WM. Drug-induced hepatotoxicity. N Engl J Med 1995;333:1118–27.

[110] Huang YS, Chern HD, Su WJ, Wu JC, Chang SC, Chiang CH, et al. Cytochrome P 450 2E1 genotype and the susceptibility to antituberculosis drug-induced hepatitis. Hepatology 2003;37:924–30.

[111] Vuilleumier N, Rossier MF, Chiappe A, Degoumois F, Dayer P, Mermillod B, et al. CYP2E1 genotype and isoniazid-induced hepatotoxicity in patients treated for latent tuberculosis. Eur J Clin Pharmacol 2006;62:423–9.

[112] Jenner AM, Timbrell JA. *In vitro* microsomal metabolism of hydrazine. Xenobiotica 1995;25:599–609.

[113] Jenner AM, Timbrell JA. Influence of inducers and inhibitors of cytochrome P450 on the hepatotoxicity of hydrazine *in vivo*. Arch Toxicol 1994;68:349–57.

[114] Desta Z, Soukhova NV, Flockhart DA. Inhibition of cytochrome P450 (CYP450) isoforms by isoniazid: potent inhibition of CYP2C19 and CYP3A. Antimicrob Agents Chemother 2001;45:382–92.

[115] Kim SH, Kim SH, Bahn JW, Kim YK, Chang YS, Shin ES, et al. Genetic polymorphisms of drug-metabolizing enzymes and anti-TB drug-induced hepatitis. Pharmacogenomics 2009;10(11):1767–79.

[116] Yimer G, Ueda N, Habtewold A, Amogne W, Suda A, Riedel KD, et al. Pharmacogenetic & pharmacokinetic biomarker for efavirenz

based ARV and rifampin based anti TB drug-induced liver injury in TB-HIV infected patients. PLoS One 2011;6:1–8.

[117] Cho HJ, Koh WJ, Ryu YJ, Ki CS, Nam MH, Kim JW, et al. Genetic polymorphisms of NAT2 and CYP2E1 associated with anti tuberculosis drug-induced hepatotoxicity in Korean patients with pulmonary tuberculosis. Tuberculosis (Edinb) 2007;87(6):551–6.

[118] Mahmoud BL, Ghozzi H, Kamoun A, Hakim A, Hachicha H, Hammami S, et al. Polymorphism of the N-acetyltransferase 2 gene as a susceptibility risk factor for anti-tuberculosis drug-induced hepatotoxicity in Tunisian patients with tuberculosis. Pathol Biol (Paris) 2012;60(5):324–50.

[119] Hiratsuka MY, Kushikawa Y, Takekuma M, Matsuura M, Narahara K, Inoue T, et al. Genotyping of the N-acetyltransferase 2 polymorphism in the prediction of adverse reactions to isoniazid in Japanese patients. Drug Metab Pharmacokinet 2002;17:357–62.

[120] Leiro-Fernandez V, Valverde D, Vazques-Gallardo R, Botana-Rial M, Constenla L, Agúndez JA, et al. N-acetyl transferase 2 polymorphisms and risk of anti-tuberculosis drug-induced hepatotoxicity in Caucasians. Int J Tuberc Lung Dis 2011;15(10):1403–8.

[121] Roy B, Chowdhury A, Kundu S, Santra A, Dey B, Chakraborty M, et al. Increased risk of anti tuberculosis drug-induced hepatotoxicity in individuals with glutathione S-transferase M1 null mutation. J Gastroenterol Hepatol 2001;16:1033–7.

[122] Yue J, Peng RX, Yang J, Komg R, Liu J. CYP2E1 mediated isoniazid-induced hepatotoxicity in rats. Acta Pharmacol Sin 2004;25:699–704.

[123] Hayashi S, Watanabe J, Kawajiri K. Genetic polymorphisms in the 5’-flanking region change transcriptional regulation of the human cytochrome P450IIE1 gene. J Biochem 1991;110(4):559–65.

[124] Lee SW, Chung LS, Huang HH, Chuang TY, Liou YH, Wu LS. NAT2 and CYP2E1 polymorphisms and susceptibility to first-line anti-tuberculosis drug-induced hepatitis. Int J Tuberc Lung Dis 2010;14(5):622–6.

[125] Roy B, Ghosh SK, Sutradhar D, Sikdar N, Mazumder S, Barman S. Predisposition of anti tuberculosis drug-induced hepatotoxicity by cytochrome P450 2E1 genotype and haplotype in pediatric patients. J Gastroenterol Hepatol 2006;21:781–6.

[126] Teixeira RL, Morato RG, Cabello PH, Muniz LM, Moreira Ada S, Kritski AL, et al. Genetic polymorphisms of NAT2, CYP2E1, and GST enzymes and the occurrence of anti-tuberculosis drug-induced hepatitis in Brazilian TB patients. Mem Inst Oswaldo Cruz 2011;106:716–24.

[127] An H, Wu X, Wang Z, Zhang JX, Liang Y. NAT2 and CYPE1 polymorphisms associated with antituberculosis drug-induced hepatotoxicity in Chinese patients. Clin Exp Pharmacol Physiol 2012;39:535–43.

[128] An H, Wu X, Wang Z, Liang Y, Zhang JX. The association of polymorphism of N-acetyl transferase 2 gene is associated with anti-tuberculosis drug-induced hepatotoxicity in tuberculosis patients. Zhonghua Yu Fang Xue Za Zhi 2011;45:36–40.

[129] Tang S, Lv X, Zhang Y, Wu SS, Yang ZR, Xia YY, et al. CYP2E1, GSTM1 and GSTT1 genetic polymorphisms and susceptibility to anti-tuberculosis drug-induced hepatotoxicity: a nested case-control study. J Clin Pharm Ther 2012;37:588–93.

[130] Strange RC, Jones PW, Fryer AA. Glutathione S-transferase: genetics and role in toxicology. Toxicol Lett 2000;112–113:357–63.

[131] Huang YS, Su WJ, Huang YH, Chen CY, Chang FY, Lin HC, et al. Genetic polymorphisms of manganese superoxide dismutase, NAD(P)H: quinine oxido reductase, glutathione S-Transferase M1 and T1 and the susceptibility to drug-induced liver injury. J Hepatol 2007;47:128–34.

[132] Chang J, Liu E, Lee C, Lin YC, Yu MC, Bai KJm, et al. UGT1A1 polymorphisms associated with risk of induced liver disorders by anti-tuberculosis medications. Int J Tuberc Lung Dis 2012;16:376–8.

[133] Laverry H, Antoine D, Benson C, Chaponda M, Williams D, Kevin Park B. The potential of cytokines as safety biomarkers for drug-induced liver injury. Eur J Clin Pharmacol 2010;66:961–76.

[134] Kim S, Kim S, Yoon H, Shin DH, Park SS, Kim YS, et al. TNF-alpha genetic polymorphism 308G/A and antituberculosis drug-induced hepatitis. Liver Int 2012;32(5):809–14.

[135] Sharma SK, Balamurugan A, Saha PK, Pandey RM, Mehra NK. Evaluation of clinical and immunogenetic risk factors for the development of hepatotoxicity during anti-tuberculosis treatment. Am J Respir Crit Care Med 2002;166:916–9.

[136] Jayaram R, Gaonkar S, Kaur P, Suresh BL, Mahesh BN, Jayashree R, et al. Pharmacokinetics and pharmacodynamics of rifampin in an aerosol infection model of tuberculosis. Antimicrob Agents Chemother 2003;47:2118–24.

[137] Estivill X, Govea N, Barcelo E, Badenas C, Romero E, Moral L, et al. Familiar progressive sensorineural deafness is mainly due to the mtDNA A1555G mutation and is enhanced by treatment with aminoglycosides. Am J Hum Genet 1998;62:27–35.

[138] Purohit P, Stern S. Interactions of a small RNA with antibiotic and RNA ligands of the 30S subunit. Nature 1994;370:659–62.

[139] Hamasaki K, Rando RR. Specific binding of aminoglycosides to a human rRNA construct based on a DNA polymorphism, which causes aminoglycoside-induced deafness. Biochemistry 1997;36:12323–8.

[140] Guan MX, Fishel-Ghodsian N, Attardi G. A Biochemical basis for the inherited susceptibility to aminoglycoside ototoxicty. Hum Mol Genet 2000;9:1787–93.

[141] Casano RA, Johnson DF, Bykhovskaya Y, Torricelli F, Bigozzi M, Fischel-Ghodsian N. Inherited susceptibility to aminoglycoside ototoxicity: genetic heterogeneity and clinical implications. Am J Otolaryngol 1999;20:151–6.

[142] Neefs JM, Van de Peer Y, De Rijk P, Goris A, De Wachter R. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res 1991;19:1987–2015.

[143] Hawkins MT, Lewis JH. Latest advances in predicting DILI in human subjects: focus on biomarkers. Expert Opin Drg Metab Toxicol 2012;8(12):1521–30.

[144] Singla R, Sharma SK, Mohan A, Makharia G, Sreenivas V, Jha B, et al. Evaluation of risk factors for antituberculosis treatment-induced hepatotoxicity. Indian J Med Res 2010;132:81–6.

[145] van der Zee AM, Alfirevic A. Pharmacogenetics and pharmacogenomics: practical applications in routine medical practice. Genome Med 2010;2:69–71.

This page is intentionally left blank

Chapter 37

Hepatitis C Virus

A.J. Thompson,[a] **P. Clark,**[b] **and A. Aghemo**[c]

[a]*Department of Gastroenterology, St Vincent's Hospital, University of Melbourne, Victorian Infectious Diseases Reference Laboratory, Australia,* [b]*Department of Gastroenterology, Princess Alexandra Hospital and School of Medicine, University of Queensland, Australia,* [c]*A.M. and A. Migliavacca Center for Liver Disease, First Division of Gastroenterology, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Università degli Studi di Milano, Italy*

Chapter Outline

37.1 Overview 859
37.2 *IL28B* Polymorphism and HCV Outcomes 860
  37.2.1 Genotype 1 HCV Treatment Outcome 860
  37.2.2 Spontaneous HCV Clearance 862
  37.2.3 *IL28B* Polymorphism and Other Clinical Phenotypes 863
  37.2.4 Mechanistic Studies and *IFN-lambda-4* *863*
37.3 *ITPA* Polymorphisms and Ribavirin-Induced Hemolytic Anemia 864
37.4 Genetic Determinants of Liver Fibrosis Progression 865
37.5 Clinical Translation 866
  37.5.1 Peginterferon and Ribavirin Therapy for Genotype 1 HCV 866
  37.5.2 Direct-Acting Antiviral Therapy for Genotype 1 HCV 869
  37.5.3 Clinical Translation of *ITPA* Genotyping 871
  37.5.4 Genomics and Drug Development 871
37.6 Conclusion 872
Acknowledgments 872
Disclosures 872
Abbreviations 872
References 873

37.1 OVERVIEW

The human genome contains about 3 billion nucleotides, which are the building blocks of DNA. Although it is mostly conserved in the population, variation in the nucleotide sequence between individuals may lead to differences in gene expression or function. During the past decade, spectacular progress in technology and bioinformatics has profoundly changed the way human genetic research can be performed. Building on the success of the Human Genome Project, in which the 3 billion nucleotides constituting our genome were first decoded [1], the international research community has continued working together to catalog the genomic variation present at the population level through the HapMap project [2]. This effort has not only resulted in a wealth of new information about human diversity, but has also represented a key step in the development of genome-wide association studies (GWASs). In a GWAS, scientists search for statistical associations between variants present in the genome and any phenotype of interest, typically disease susceptibility, but more recently drug response and toxicity, hoping to identify changes that can help personalize patient care and lead to a better understanding of underlying pathophysiological mechanisms.

The most common form of genetic variation in the human genome is the single-nucleotide polymorphism (SNP). There are more than 10 million common SNPs in any given population, where a "common SNP" has a minor allele frequency of ≥5%. The HapMap project defined patterns of linkage disequilibrium (LD) between SNPs. LD describes the tendency of neighboring variants to be inherited together as part of genomic blocks called haplotypes (Figure 37.1). This allows SNPs to be identified that "tag" haplotype blocks and indirectly represent multiple other SNPs (tag SNPs). The SNPs genotyped on a GWAS chip are selected to represent haplotypes across the entire genome, covering common variants for all genes, and they contain information about almost all common variants present in human populations. The genome-wide genotyping chips used in GWASs contain hundreds of thousands—now up to 5 million—tag SNPs.

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00037-2
© 2014 Elsevier Inc. All rights reserved.

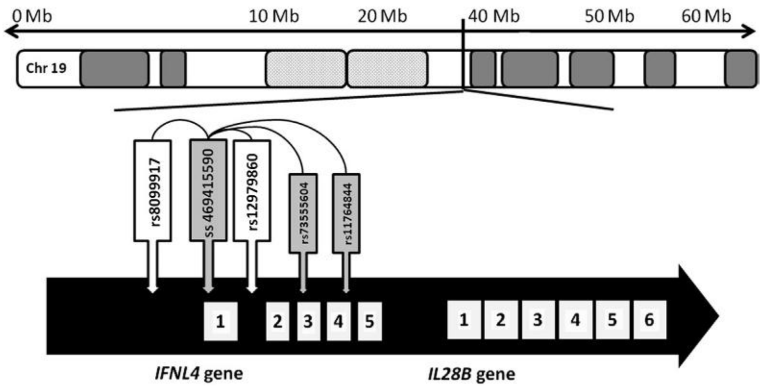


FIGURE 37.1 **GWAS use of tag SNPs to identify loci associated with a phenotype of interest.** rs12979860 and rs8099917 were the two SNPs identified in the original discovery studies that tag a haplotype in the region of the *IL28B* gene on chromosome 19. Recent data suggest that they may in fact tag the newly identified *IFNL4* gene, which is upstream of *IL28B* [44]. ss469415590 is in strong linkage disequilibrium with rs12979860. Two additional nonsynonymous variants in *IFNL4* have subsequently been identified: rs73555604 (p.Cys17Tyr), and rs11764844 (p.Pro70Ser), which may modify risk in carriers of the unfavorable ss469415590[ΔG] [44].

It is this ability of tag SNPs to indirectly represent non-genotyped SNPs that makes the technology truly genome-wide and particularly attractive for non-a-priori discovery genetic studies. This is the scientific approach that was used to identify the association between genetic variation in the *IL28B* gene region and peginterferon-α (pegIFN) and ribavirin (RBV) treatment response in patients chronically infected with the hepatitis C virus (HCV), as well as spontaneous clearance of acute HCV infection.

37.2 *IL28B* POLYMORPHISM AND HCV OUTCOMES

37.2.1 Genotype 1 HCV Treatment Outcome

Four genome-wide association studies published in late 2009 and early 2010 independently identified a strong association between genetic variation in the region of the *IL28B* gene and pegIFN and RBV treatment response in patients chronically infected with genotype 1 HCV (HCV-1).

The first genome-wide analysis of common genetic variants associated with HCV-1 outcome interrogated 1604 patients from the IDEAL study. This was a large phase 4 comparative effectiveness study of the two commercially available pegIFN preparations (pegIFNα-2b versus pegIFNα-2a) in North American patients chronically infected with HCV-1 [3]. A further 67 patients were enrolled from a second randomized clinical trial that evaluated the efficacy of pegIFNα-2b plus RBV in Caucasians versus African Americans [4]. Samples were genotyped using the Illumina Human610-Quad BeadChip (Illumina, San Diego, CA). The primary analysis for treatment outcome compared sustained virological response (SVR) to true biological nonresponse in 1137 patients of European American (n = 971), African American (n = 191), or Hispanic (n = 75) ancestry. Biological nonresponse was defined as no SVR despite adherence ≥80%. Seven SNPs on chromosome 19 in the region of the *IL28B* gene, coding for IFN-λ-3, met the threshold for Bonferroni-corrected genome-wide significance [3].

The top discovery SNP was rs12979860, located 3 kb upstream of the *IL28B* gene and having a p-value for association of 1.06×10^{-25} in European Americans ($p = 1.37 \times 10^{-28}$ combined across the three ethnic populations). This is a bi-allelic polymorphism with three possible genotypes (C/C, C/T, and T/T). In this adherent cohort, the good-response C/C *IL28B* variant was associated with a 2-fold increase in SVR rate in all ethnic groups. Although the heterozygote C/T variant was associated with a small increase in SVR rate compared to the T/T variant, the difference was not clinically meaningful. Six other SNPs exceeded the statistical threshold for genome-wide significance, including the polymorphism rs8099917. All displayed different degrees of LD with rs12979860, and their effects were largely explained by this polymorphism. None of these seven associated SNPs were obviously functional.

A number of other important observations were made in this cohort. The *IL28B* polymorphism was associated with early viral decline on treatment, with higher rates of viral clearance at week 2, week 4, and week 12 in patients with the good-response *IL28B* genotype (Figure 37.2) [5]. This has since been shown to be a dominant effect on phase 1 kinetics, with clear differences in viral load reduction as early as 24 hours following the first injection of pegIFN [6]. The favorable C/C genotype was more common in European Americans compared to African Americans, and it was estimated that this difference in genotype frequency was responsible for more than half of the discrepancy in treatment response rates between the two populations.

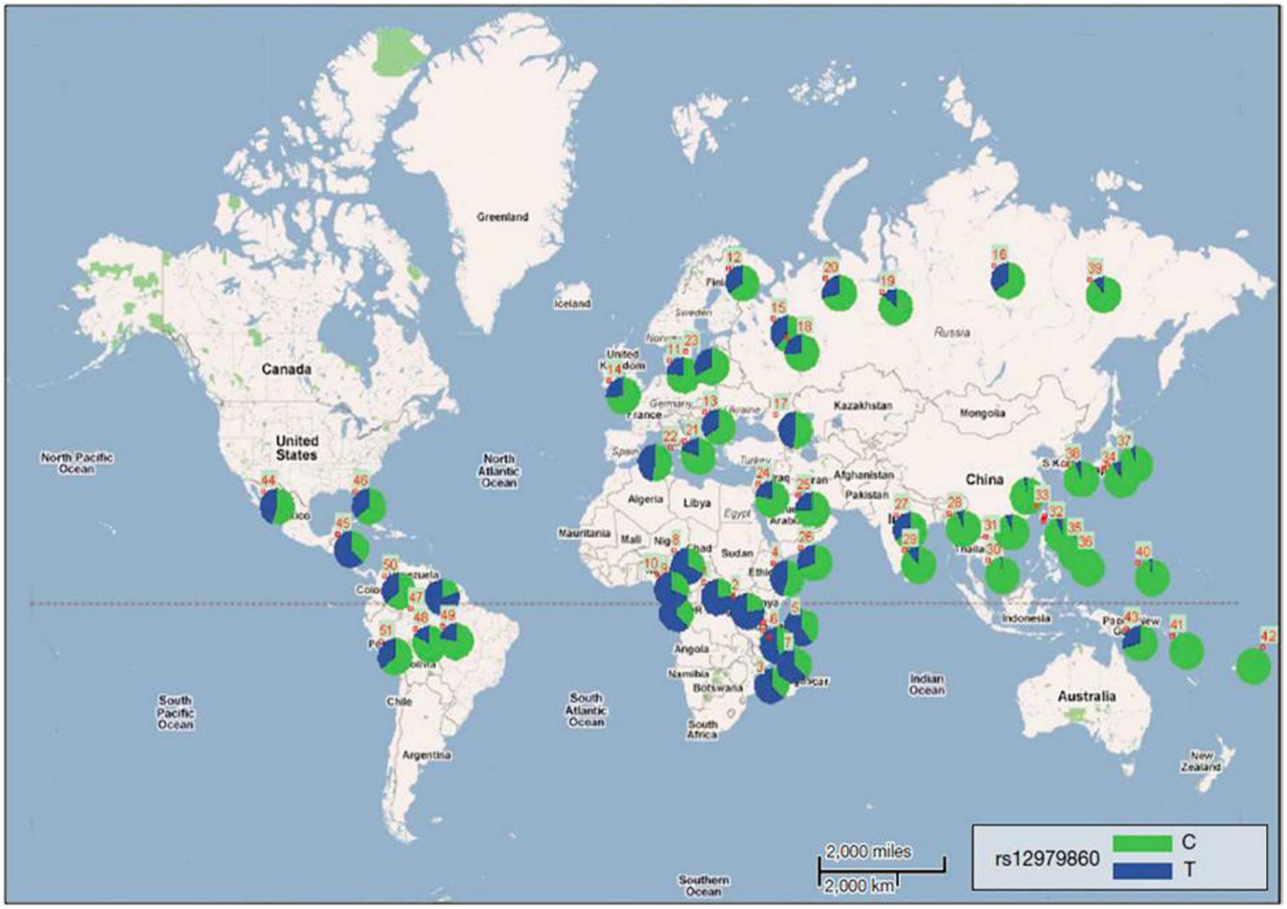

Greenland
Iceland
Canada
United States
North Pacific Ocean
North Atlantic Ocean
South Pacific Ocean
South Atlantic Ocean
Russia
Kazakhstan
Mongolia
China
Afghanistan
Pakistan
Algeria
Libya
Egypt
Mauritania
Mali
Chad
Sudan
Angola
Namibia
Botswana
South Africa
Chile
Argentina
Indonesia
Indian Ocean
Australia
New Zealand
Southern Ocean
2,000 miles
2,000 km
rs12979860
C
T

FIGURE 37.2 Data from Caucasian patients treated with pegIFN and RBV in the IDEAL pharmacogenomics cohort (intent-to-treat analysis) [5]. (a) Virological response rates according to *IL28B* genotype (rs12979860). This genotype is a *pretreatment* predictor of on-treatment virological milestones, including RVR and cEVR, as well as SVR. (b) SVR rates according to on-treatment response and *IL28B* genotype. RVR and cEVR remain important *on-treatment* predictors of IFN treatment outcomes. The *IL28B* genotype frequency for each responder group is shown in the pie chart. RVR4—rapid virological response, undetectable HCV RNA at week 4; cEVR—complete early virological response, undetectable serum HCV RNA at week 12; pEVR—partial EVR, >2 log reduction in serum HCV RNA from baseline to week 12 but persistent viraemia. *Non-RVR4—cEVR + pEVR + NR.

In fact, the frequency of the protective allele is now recognized to be highly divergent between ethnic groups, with Asians > Caucasians (Northwestern Europe > North America) > Hispanics > Africans [7,8] (Figure 37.3). This is generally consistent with the pegIFN and RBV response rates that have been observed in different populations [4,9].

The *IL28B* polymorphism was also implicated in spontaneous clearance of HCV infection. The C allele frequency in the study population was lower than in healthy controls, suggesting selection away from the protective allele. Finally, a genome-wide analysis of pretreatment viral load set point identified rs12979860 as the only common variant associated with HCV viral load, with the good-response genotype associated with a higher viral load [7]. This apparently paradoxical association may be explained by patterns of intrahepatic IFN-stimulated gene (ISG) expression, to be discussed.

A second GWAS conducted in HCV-1-infected patients of European ancestry (Australian/Western Europe) also identified an important role for the *IL28B* gene region in determining HCV treatment response [10]. Investigators tested for polymorphisms associated with SVR in a clinic population. Adherence to therapy was not specifically defined. A 2-stage approach was taken (discovery, $n = 293$/replication, $n = 555$). The discovery GWAS was performed using the Illumina CNV 370-Quad BeadChip or the Illumina Infinium HumanHapMap300. A total of 172 SNPs were taken forward for replication, selected on the basis of a p-value for association $<10^{-5}$, or $p < 10^{-3}$, and immune-regulatory/antiviral function. The polymorphism rs8099917 exceeded the threshold for genome-wide significance in the analysis of the combined cohort (OR 1.98 [1.57–2.52], $p = 9.25 \times 10^{-9}$). rs12979860 was not included on the genotyping chip and was not tested.

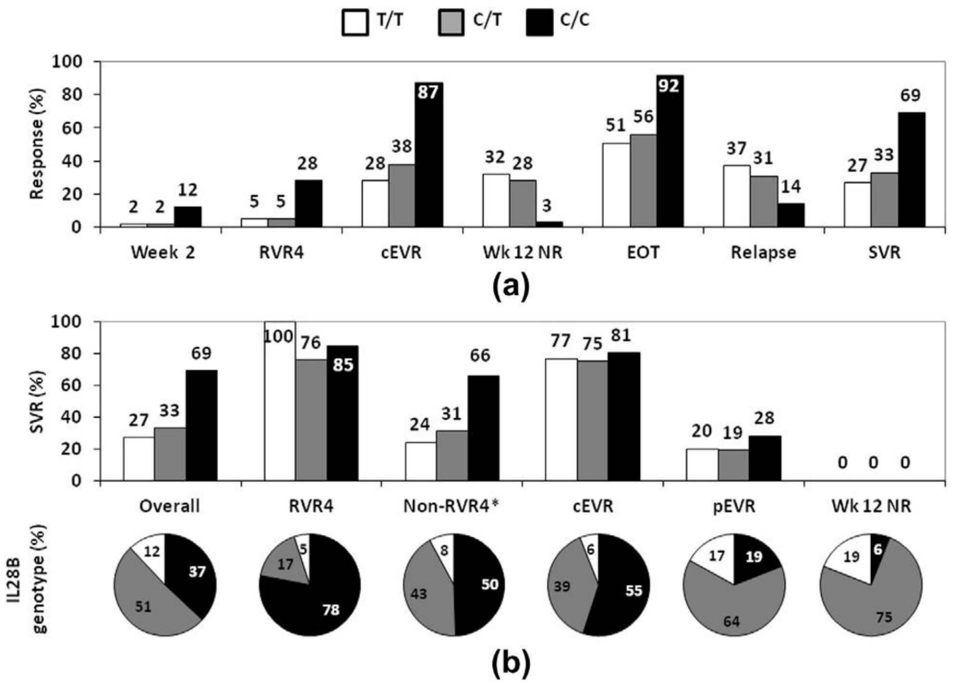


FIGURE 37.3 World map of rs12979860 allele frequency. *Source: From Thomas et al.* [8].

Japanese investigators conducted a third GWAS in an HCV-1B population [11]. The 2-stage design included a discovery cohort of 154 Japanese patients and an ethnically matched replication cohort of 172 patients. The Affymetrix SNP 6.0 genome-wide SNP typing array (Affymetrix, Santa Clara, CA) was used. The analysis tested for variants associated with null virological response (NVR) (defined as serum HCV RNA reduction <2log$_{10}$ IU/mL at week 12 of therapy and detectable HCV RNA at week 24). Adherence >80% during the first 12 weeks of treatment was required for inclusion. Two SNPs in the *IL28B* gene region were strongly associated with NVR, rs8099917, and rs12980275 (combined $p = 2.68 \times 10^{-32}$, OR 27.1 [14.6–50.3] and $p = 2.84 \times 10^{-27}$, OR 17.7 [10.0–31.3], respectively). Again, the tag SNP rs12979860 was not genotyped. The large ORs observed can be explained by the extreme phenotype—virtually no patients homozygous for the good-response allele qualified for NVR (also noted in the IDEAL cohort, where <5% of patients with the good-response genotype were NVR) [7,11]. In a secondary analysis, the SNPs rs8099917 and rs12980275 were confirmed to be associated with SVR versus no SVR.

Finally, the results of a fourth GWAS, by Rauch and colleagues, came from the Swiss Hepatitis C and HIV Cohort Studies. Although the primary aim of this study was to analyze genetic predictors of spontaneous clearance (see Section 37.2.2), the control cohort of patients with persistent HCV infection included 465 individuals for whom pegIFN and RBV treatment response data were available. All patients were ≥80% adherent to therapy. Patients were infected with HCV-1/4 (n = 188/34) or HCV-2/3 (n = 60/162). Genotyping was performed using the Illumina Human1M-Duo, HumanHap550, or Human610W-Quad BeadChips. *IL28B* genotype was significantly associated with treatment response (rs8099917, $p = 3.11 \times 10^{-8}$).

Multiple retrospective candidate gene studies have since replicated the observation that a polymorphism within a common haplotype in the region of the *IL28B* gene is a key predictor of treatment response in genotype 1 HCV [12–21].

37.2.2 Spontaneous HCV Clearance

The *IL28B* polymorphism is strongly associated with spontaneous clearance of acute HCV infection. Thomas and colleagues evaluated *IL28B* genotype (rs12979860) frequency in a large cohort of individuals with either spontaneous clearance of HCV (n = 388) or persistent HCV infection (n = 620). The frequency of the good-response C/C genotype was strongly associated with spontaneous clearance in both white and black patients (OR = 3.0 [2.2–4.0], $p = 3 \times 10^{-13}$). The effect of the locus on spontaneous clearance was similar between patients of European and African ancestry, and was not altered by HBV or HIV coinfection. A genome-wide association study of spontaneous HCV clearance subsequently provided independent validation of this association (rs8099917, OR = 2.3 [1.7–3.1], $p = 6.1 \times 10^{-9}$). *IL28B* variants were the only common variants associated with natural clearance [22]. The association between *IL28B* genotype and spontaneous clearance has now been confirmed in multiple replication cohorts [23–25].

37.2.3 *IL28B* Polymorphism and Other Clinical Phenotypes

Studies have investigated the association between the *IL28B* polymorphism and a number of other clinical and laboratory phenotypes, including liver histology, necro-inflammatory activity, fibrosis, and steatosis, as well as serum lipid levels.

Further exploration of the IDEAL pharmacogenomics cohort identified *IL28B* polymorphism to be associated at the genome-wide level with serum ALT level as well as necro-inflammatory activity on liver biopsy, where the good-response variant was associated with a higher-level serum ALT as well as higher-grade activity [26]. No other common variants were associated with either ALT level or hepatic necro-inflammation. There was no significant association between the *IL28B* genotype and METAVIR fibrosis stage in the same cohort [27]. Similarly, there was no significant association noted between *IL28B* genotype and fibrosis stage in a second cross-sectional study of 496 Japanese patients [28]. No relationship between *IL28B* genotype and fibrosis progression rate over time was noted in an Italian study of 297 patients in longitudinal follow-up [29]. However, A second study from Japan reported a positive correlation between the good-response variant and both histological activity and liver fibrosis stage (F2-4) [30]. The good-response *IL28B* genotype (rs12979860) was also associated with more prevalent cirrhosis, as defined by Fibroscan >15.5 kPa, in a Spanish cohort of patients with HCV-HIV coinfection [31]. In other series, the poor-response *IL28B* genotypes was linked to advanced liver fibrosis [32,33], although this may reflect selection bias for patients who have failed antiviral therapy.

As can be seen, the existing data, mostly cross-sectional, that link *IL28B* genotype to liver fibrosis stage are conflicting. There is a need for more prospective studies that test for genetic associations with fibrosis progression rate.

Chronic infection with HCV has been associated with low serum lipid levels and hepatic steatosis, both of which correct after viral eradication. The *IL28B* polymorphism was recently associated with both dyslipidemia and hepatic steatosis in the setting of HCV-1 infection. Serum total cholesterol (TC), low-density lipoprotein (LDL), and apolipoprotein B levels have been reported to be significantly lower in patients with the poor-response *IL28B* genotypes [34,35]. Interestingly, this relationship is no longer present following viral eradication [36]. The poor response *IL28B* genotypes have also been associated with greater degrees of hepatic steatosis in HCV [35,37]. This might suggest a common mechanism linking *IL28B*, lipid metabolism, and hepatic steatosis. The researchers hypothesized that these associations might be explained by levels of intrahepatic ISG expression (see later), consistent with the known effect of exogenous IFN to reduce serum LDL levels [38].

37.2.4 Mechanistic Studies and *IFN-lambda-4*

The mechanism by which *IL28B* variants influence IFN responsiveness was not identified by the discovery GWASs. The studies generated a number of key questions, among them: What is the causal variant, and how does this mediate the downstream immunological phenotype?

Initial attempts to identify a causal variant focused on pure genetic approaches, using association statistics in the discovery studies. They were largely unsuccessful because of the high linkage disequilibrium in the region [7,39]. Putative causal variants were identified by fine mapping/resequencing. These included a nonsynonymous coding variant in exon 2 of *IL28B* (rs8103142), a variant near the translation initiation codon (rs28416813, G>C), two promoter SNPs (rs4803219, C>T; rs28416813, C>G) and two 3′-UTR SNPs (rs4803217, C>A; rs581930, C>T). All are in very high LD with the top discovery tag SNPs [7,39]. Functional studies of the most likely candidate, rs8103142, have not confirmed any differential role in *IL28B* innate immune stimulation or antiviral efficacy [40]. Differential expression of *IL28B* has not been demonstrated in stimulated hepatocytes, infected hepatocytes [40–43], stimulated immune cells, or immune cells of individuals infected with HCV [43].

Very recently, it has been proposed that the GWAS discovery polymorphisms may in fact tag a new gene, *interferon-lambda-4* (*IFNL4*) [44]. A dinucleotide polymorphism (ss469415590, ΔG/TT) has been identified upstream of *IL28B* that creates (ΔG) or disrupts (TT) an open reading frame in *IFNL4*. Individuals who carry the ΔG allele can generate the IFNL4 protein. IFNL4 is not produced by individuals who are homozygous for the TT allele because of a frameshift in exon 1 caused by the insertion variant. ss469415590 is in strong linkage disequilibrium (LD) with the previously identified tag SNPs in individuals of European or Asian ancestry—rs12979860 is actually located within intron 1 of *IFNL4* (Figure 37.1) [44].

In contrast, linkage disequilibrium is only moderate in individuals of African ancestry [44]. This weaker LD has been exploited to show that, in African American individuals, ss469415590 appears to be more closely associated with IFN responsiveness as well as with spontaneous clearance of HCV, than is rs12979860 [44,45]. The ΔG *IFNL4* allele was more closely associated with outcome than the poor-response T allele of rs12979860. The data are consistent with a critical role for IFNL4 in regulating exogenous IFN response, with the simple hypothesis that ss469415590 determines the expression of IFNL4, and that IFNL4 negatively regulates type 1 IFN signaling.

Rarely is biology simple, however. Haplotype analysis has shown that, although the ss469415590 TT allele is present on a unique favorable haplotype, the unfavorable ΔG allele occurs on a number of haplotypes, including two that were previously reported to be neutral in Europeans despite carrying the unfavorable rs12979860 T allele [46,47]. If ss469415590 alone explains IFN responsiveness, all haplotypes carrying the ΔG allele should have equal risk. The genetics may therefore be more complicated.

Two nonsynonymous variants in *IFNL4* have subsequently been identified—rs73555604 (p.Cys17Tyr) and rs11764844 (p.Pro70Ser)—that may mitigate the effect of the IFNL4 protein [44]. Discordant haplotypes were noted to include minor alleles of either variant [44]. This requires further investigation. Although the data are compelling, we refer to the *IL28B* polymorphism rather than the *IFNL4* polymorphism in the rest of the chapter, given that the existing literature has been generated using *IL28B* genotyping.

IL28B and IFNL4 are both members of the type 3 IFN family [48], also known as the IFN-lambda family (including IL29 (IFNL1), IL28A (IFNL-2) and *IL28B* (IFNL3), and IFNL4). Expression of type 3 IFNs is triggered by most stimuli that induce type 1 IFN, and both cytokine families activate similar downstream signaling pathways to induce ISGs. Type 3 IFNs inhibit viral replication *in vitro* and *in vivo*, and recombinant IL29 is currently in phase 2/3 clinical development for the treatment of both HCV and HBV [49]. A key point of distinction from the type 1 IFN axis is that IFNL1–3 signal via the unique type 3 IFN receptor. In contrast to the ubiquitous type 1 IFN receptor, the distribution of the type 3 IFN receptor appears to be tissue-dependent; it is expressed at a high level in liver and epithelial cells but at a low level in hemopoietic cells and in the central nervous system. Type 3 IFN inhibits HCV replication in cell culture models [16,17] as well as *in vivo* [18]. Recent data suggest that type 3 IFN may be the dominant IFN produced by the liver in response to HCV infection [19,20].

IFNL4 is the newest member of the type 3 IFN family. Although it is more similar to the members of the IFN-lambda family (IL28A, IL28B, and IL29, or IFNL1–3) than is any other protein, it shares only 29% amino acid identity and is therefore very different [44]. In primary human hepatocytes, IFNL4 is transiently transcribed and translated following stimulation with poly (I:C), an analog of viral double-stranded RNA. Expression of IFNL4 in HepG2 cells leads to induction of ISGs via the Janus kinase-signal transducers and activators of the transcription (JAK-STAT) pathway. Somewhat surprisingly, treatment of cells with recombinant IFNL4 does not stimulate the JAK-STAT pathway, suggesting that this pathway is activated by cytoplasmic expression intracellularly. This makes it distinct from the other members of the type 3 IFN family, which signal via the external type 3 IFN receptor. Again, further functional studies are required in this very novel and fertile area.

Regardless of the exact genetic mechanism, it is likely that intrahepatic ISG expression is central to the biology involved. The favorable *IL28B* genotype has been associated with low-level ISG expression in the liver [40–42], consistent with the historical observation that low levels of liver ISG expression predict for rapid IFN treatment response [50]. Conversely, unfavorable *IL28B* genotypes are associated with higher-level intrahepatic ISG expression and poor responsiveness to exogenous pegIFN therapy. Low-level intrahepatic ISG expression might also explain the observation that the favorable *IL28B* variant is associated with a higher viral load set point [7]. Differential expression of intrahepatic ISGs may now be explained by the presence, or absence, of IFNL4 expression; this needs to be confirmed *in vivo*. The *IL28B* genotype has also been associated with NK cell activity [26] as well as with peripheral blood mononuclear cell markers of innate immunity [43]. Further investigation of the relationship between the *IFNL4*/*IL28B* genotype, hepatic ISG expression, and immunological phenotype is therefore required.

We are only just starting to understand the biological mechanism for the *IL28B* association. The innate immune system is implicated by the strong association with phase 1 kinetics, the intrahepatic ISG data, and, as discussed later, the observed association between donor liver IL28B genotype and IFN response in the post–liver transplant setting. Detailed *ex vivo* and *in vitro* experiments will be required, opening up a new field of *IL28B* immunovirology.

37.3 *ITPA* POLYMORPHISMS AND RIBAVIRIN-INDUCED HEMOLYTIC ANEMIA

Two functional polymorphisms in the *ITPA* gene on chromosome 20 have more recently been associated with ribavirin-induced hemolytic anemia in a separate GWAS performed using the IDEAL pharmacogenetics cohort [51]. Several tag SNPs on chromosome 20 were associated with quantitative reduction in hemoglobin (Hb) at week 4 of treatment. The top discovery SNP was rs6051702 (patients of European American ancestry, p-value = 1.1×10^{-45}) [51]. The *ITPA* polymorphism was also associated with Hb reduction >3 g/dL and Hb level <10 g/dL at week 4.

The genetic locus identified contains the *ITPA* gene, coding for the enzyme inosine triphosphatase (ITPase). Functional *ITPA* polymorphisms were previously identified that result in reduced ITPase activity [52–55] and accumulation of inosine triphosphate (ITP) in red blood cells, including rs1127354 and rs7270101. Rs1127354 is a missense variant in exon 2 of the *ITPA* gene (P32T); rs7270101 is a splicing-altering variant located in the second intron (IVS2). Neither were included on the genotyping chip, but they were noted to be in LD with rs6051702 using data from HapMap, where the low-activity variants were associated with the protective rs6051702 C allele. The cohort was then genotyped for the two functional variants. The two variants cosegregated with the top association variant, each was independently associated with anemia, and the two were shown to entirely explain the discovery association signal [51]. The p-value for association of a combined low-activity *ITPA* allele made up of either functional variant was $p = 2.23 \times 10^{-91}$. On the basis of this combined allele, 14% of the study population ($n = 186$) were predicted to have severe ITPase deficiency (<30% ITPase activity). None dropped Hb to a level <10 g/dL at week 4; only 4% experienced Hb reduction >3 g/dL (compared respectively to 12% and 56% of patients with wild-type activity) [51]. The association between *ITPA* polymorphism and treatment-induced anemia has since been replicated in a number of independent cohorts [15,56–59]. The association is independent of HCV genotype [57].

ITPase is the enzyme responsible for the hydrolysis of the inosine triphosphate (ITP) protein. ITPase deficiency was originally described as a benign red cell enzymopathy. Functional variants result in reduced ITPase activity and lead to accumulation of ITP in erythrocytes. A mechanism for the protective effect of the *ITPA* genotype was recently proposed. RBV depletes erythrocyte guanosine triphosphate (GTP), leading to ATP depletion, oxidative stress, and hemolysis. ITP is able to substitute for GTP in the biosynthesis of ATP, thereby protecting against RBV-hemolysis [60].

More recently the *ITPA* polymorphism was also associated with pegIFN-induced thrombocytopenia, with the effect in the opposite direction (low-activity variants predict for more severe thrombocytopenia) [59,61]. In fact, this an indirect genetic association, where normal ITPase activity is associated with more severe hemolytic anemia, driving a reactive thrombocytosis that manifests as a relative amelioration of pegIFN-related thrombocytopenia. Therefore, in patients with ITPase deficiency who are protected from RBV-hemolysis, more severe thrombocytopenia is observed. Genome-wide analysis for predictors of pegIFN-related leukopenias in the IDEAL cohort did not identify any significant associations [61].

37.4 GENETIC DETERMINANTS OF LIVER FIBROSIS PROGRESSION

The discovery that *IL28B*/*ITPA* variants act as predictors of HCV treatment outcomes has generated much hope that genome-wide screening might lead to similar insights into the understanding of liver fibrosis progression.

The identification of factors associated with disease progression in chronic hepatitis C has been a major focus in the field over the last two decades. Indeed, the availability of a treatment with limited efficacy and poor tolerability (e.g., pegIFN and RBV) called for the identification of patients who either required treatment immediately or could be safely deferred until the arrival of more effective and/or better tolerated regimens. Much emphasis has been put on the important role of exogenous factors such as alcohol consumption, HIV coinfection, or metabolic diseases (e.g., diabetes, insulin resistance) on the rate of fibrosis progression [62]. However, these factors are estimated to contribute no more than 30% of the individual variability observed in rates of fibrosis progression.

This finding indirectly supports a role for host genetics in determining fibrosis progression in chronic hepatitis C. Although progression rate is a difficult phenotype to accurately define—given the limitations of histological stage as a gold standard reference and the difficulties involved in the precise characterization of disease duration, as well as important confounders such as alcohol consumption—this observation indirectly supports a role for host genetics in determining prognosis.

Most early studies used a relatively straightforward candidate gene approach, in which cohorts were analyzed for the allele frequency of SNPs affecting plausible genes according to the presence or absence of advanced liver fibrosis. Unfortunately, the candidate gene literature has been generally limited by small cohort size and the absence of independent replication, making it difficult to draw firm conclusions. The data are therefore difficult to evaluate. Candidate genes identified include *matrix metalloproteinases (MMP-1/3/9)* and *monocyte chemotactic protein 1 (MCP-1)* [63,64].

The CCA haplotype in the vitamin D receptor gene was identified as a risk marker for fibrosis progression in a third study that used a more stringent definition for fibrosis progression: evaluation of paired biopsies from the same patient performed after a congruous period of time in the absence of effective anti-HCV treatments [65]. However, the most promising genetic variant to emerge from candidate gene studies is a SNP in the gene known as *patatin-like phospholipase-3 (PNPLA3)*. *PNPLA3* encodes a 481–amino acid protein of unknown function that belongs to the patatin-like phospholipase family. Patatin, the progenitor of this family, has nonspecific lipid acyl hydrolase activity. A nonsynonymous polymorphism at position 148 (rs738409 C>G) encodes an isoleucine-to-methionine substitution, which has been shown to be the strongest genetic determinant of human steatosis in GWASs of nonalcoholic fatty liver disease (NAFLD) [66,67]. The G allele of rs738409 SNP was subsequently associated with the presence of histologically determined liver steatosis, cirrhosis, and accelerated fibrosis progression in patients with chronic hepatitis C infection [68,69].

More comprehensive genetic studies are now starting to emerge. In a gene-centric disease association study of 24,832 putative functional SNPs, Huang and colleagues identified a 7-SNP variant signature that was associated with the risk of advanced fibrosis in chronic HCV infection as well as with estimated fibrosis progression rate [70]. By analyzing SNPs in seven genes (*adaptor-related protein complex 3 S2*, *aquaporin 2*, *antizyme inhibitor 1*, *degenerative spermatocyte homolog 1 lipid desaturase*, *syntaxin binding protein 5-like*, *toll-like receptor 4* (*TLR4*), and *transient receptor potential cation channel M5*), the researchers created the Cirrhosis Risk Score (CRS) to estimate the risk of developing cirrhosis at the individual patient level. Three cut-offs for CRS were proposed: <0.5 low risk, 0.5–0.7 moderate risk, and >0.7 high risk. The high-risk CRS cut-off was able to correctly identify 82% of HCV patients progressing to cirrhosis as well as 90% of those with mild fibrosis at the initial liver biopsy who showed fibrosis progression at a second liver biopsy. The CRS has been independently validated in small longitudinal cohorts for predicting patients with and without progression [71,72]; it has also been validated in the setting of post–liver transplant HCV recurrence [73]. However, mechanistic data linking these SNPs to pathways of liver fibrogenesis are limited, although TLR4 has been identified as an important mediator. Further validation in large prospective cohorts is required before the CRS can be recommended as a clinical tool.

A more traditional genome-wide approach was recently used to investigate fibrosis progression. Patin and colleagues conducted a GWAS using 1161 well-characterized HCV-infected patients of European descent who had liver biopsy prior to treatment [74]. Fibrosis progression was defined categorically by METAVIR stage, as well as continuously by progression rate (stage ÷ estimated duration of infection). The researchers tested for associations with 780,650 SNPs; the top discovery SNPs (p-values $<5\times10^{-5}$) were then validated in an independent replication cohort of 962 patients. Two SNPs were significantly associated with fibrosis progression, one located in *RNF7* ($p=8\times9\times10^{-9}$) and another linked to *MERTK* ($p=2\times1\times10^{-9}$) [104]. Both are plausibly related to fibrogenesis. *RNF7* acts as an antioxidant, protecting against apoptosis, and *MERTK* is involved in phagocytosis of apoptotic cells by macrophages. However, mechanistic data are not yet available.

In conclusion, liver fibrosis progression remains an important clinical issue, and genetic predictors could prove helpful in prioritizing patients for lifestyle modification or therapeutic intervention. Although the data support a role for host genetic background as a modifier of the natural course of HCV infection, clinical utility has yet to be translated. More important, the advent of more highly effective DAA therapies in the near future is likely to lessen the impact of future studies on this topic, as universal treatment of HCV patients is likely to be recommended.

37.5 CLINICAL TRANSLATION

The *IL28B* polymorphism strongly predicts for the clinical outcome of pegIFN and RBV-based treatment of genotype 1 HCV infection. However, the combination of pegIFN and RBV is no longer the standard of care in many parts of the world. In 2011, the first direct-acting antivirals (DAAs) were approved for the treatment of HCV-1. Telaprevir and boceprevir are HCV NS3 protease inhibitors that increase SVR rates almost 2-fold [75–78]. The year 2011 also saw proof of concept that HCV could be cured using IFN-free regimens. DAA therapy clearly attenuates the association between IL28B genotype and HCV treatment response, but emerging data suggest that IL28B genotype will remain relevant to emerging treatment paradigms, including IFN-free therapy. This is an extremely dynamic area, but in the global context it is still germane to discuss the clinical utility of *IL28B* genotyping for pretreatment counseling in the context of pegIFN and RBV, as well as for the rapidly evolving DAA therapies for HCV.

37.5.1 Peginterferon and Ribavirin Therapy for Genotype 1 HCV

37.5.1.1 IL28B Genotyping

Most of the clinical studies that validated the association between IL28B genotype and pegIFN and RBV treatment response have tested one of two SNPs—either rs12979860 or rs8099917. As stated, these two polymorphisms tag a common haplotype [7] and are in strong linkage disequilibrium (LD) in Caucasians and Asians. In most patients the genotypes will be concordant. Therefore, they provide very similar information for clinicians; both predict for treatment response/nonresponse with reasonable, but not perfect, accuracy. The exception is patients of African ancestry, where rs12979860 and rs8099917 are in weaker LD, and rs12979860 is more closely associated with IFN outcome [7]. Although data have been presented showing that the accuracy of prediction of clinical outcomes may be increased by testing both SNPs [47], this requires validation in different ethnic populations, and in practice the clinical utility is uncertain. The recent identification of the *IFNL4* dinucleotide polymorphism (ss469415590) as a proposed causal variant suggests that this will be the most appropriate SNP to genotype in all populations in the future [44].

37.5.1.2 IL28B Genotype and Peginterferon/Ribavirin Therapy in HCV-1

Intent-to-treat analysis of the IDEAL pharmacogenomics cohort showed that amongst Caucasians HCV-1 patients treated with pegIFN and RBV, the rate of SVR was 69% in patients carrying the good-response *IL28B* variant versus 33% and 27% in the setting of poor response variants [5] (Figure 37.2). In regions where DAAs are anticipated, this information may be useful for informing the decision to treat immediately, or to wait for the availability of direct-acting agents (DAA) for HCV-1 [5]. Patients carrying the poor response *IL28B* genotypes have a low SVR rate with the current standard-of-care and are most suitable for deferral. The negative predictive power of *IL28B* genotype is only ~70% however [5], and it should therefore not be used alone to exclude patients from treatment, particularly in cirrhotic patients from regions where the availability of DAA-therapy is likely to be delayed. *IL28B* genotyping is now available as a commercial diagnostic or as a research test in most developed countries.

37.5.1.3 IL28B Genotyping versus On-Treatment Kinetics for Predicting SVR

IL28B genotyping and rapid virological response (RVR) are complementary, and both are associated with IFN sensitivity and viral kinetics. IL28B genotype is the strongest *baseline* predictor of virological response to treatment and has the advantage of identifying patients who are IFN responsive *before* treatment initiation. It profoundly influences phase 1 viral kinetics such that the good-response variant is associated with greater rates of RVR, complete early virological response (cEVR), and SVR. The positive predictive value (PPV) for SVR of the good-response *IL28B* genotype was 69% in the intent-to-treat analysis of Caucasian members of the IDEAL pharmacogenetics cohort [5].

In contrast, RVR is an on-treatment measure of IFN responsiveness. It selects a smaller number of patients, but its PPV is higher. In the same intent-to-treat analysis, the overall rate of RVR in the Caucasian group was 14%, but 84% of these patients attained SVR [5]. Once RVR has been achieved, *IL28B* genotype is no longer predictive of treatment outcome, and rates of SVR are high in all patients (SVR in Caucasians: rs12979860 C/C = 85%; C/T = 76%; T/T = 100%; p = NS [5]). Note that in the IDEAL study, >75% of these RVR patients carried the C/C good-response variant. Importantly, in non-RVR patients, homozygosity for the good-response variant is still associated with a 2-fold increase in SVR rate (SVR in Caucasian non-RVR patients: C/C = 66%; C/T = 31%; T/T = 24%; $p < 0.001$) [5] (Figure 37.2).

37.5.1.4 IL28B Polymorphism and Short-Treatment Duration

There is now considerable data supporting a treatment duration of 24 weeks as sufficient for noncirrhotic RVR patients with a baseline HCV RNA level <400,000–800,000 IU/mL [79]. Therefore, patients can already be treated for 24 weeks in the setting of RVR and low baseline viral load. A key issue is whether it is possible to shorten treatment in *all* good-response *IL28B* patients, including those with non-RVR or high baseline viral load. This is of interest given the high SVR rates observed in non-RVR patients in the IDEAL pharmacogenomics cohort [5].

No prospective data have compared the efficacy of 24 versus 48 weeks of pegIFN and RBV therapy for good-response *IL28B* patients. It is unlikely that studies will be performed to address this issue given the development of DAAs. However, recent data from a large retrospective series in Taiwan showed that short-duration therapy (24 weeks) is sufficient only for patients who carry the good-response *IL28B* genotype, who have a low baseline viral load, and who achieve an RVR (SVR rates 95 versus 99% with 24 versus 48 weeks) [80]. Among patients carrying the good-response *IL28B* genotype, SVR rates were significantly lower with 24 versus 48 weeks total treatment duration in the following setting: RVR plus high baseline viral load (SVR 70% versus 97%), non-RVR plus (26% versus 60%) [80].

37.5.1.5 IL28B Genotyping and PegIFN Therapy in Other Settings

*Non-genotype 1 HCV

Peginterferon and ribavirin remains the standard-of-care treatment of non-genotype 1 HCV. There is emerging evidence concerning the role of *IL28B* genotype in predicting non-genotype 1 HCV treatment response. The data for genotype 2/3 (HCV-2/3) are mixed. Compared to HCV-1, HCV-2/3 are more responsive to IFN-based therapy, requiring only a 24-week duration and achieving SVR rates >70% [79]. Rauch and colleagues included 222 adherent patients with genotype 2/3 HCV in their treatment response subanalysis. *IL28B* polymorphism was not significantly associated with treatment outcome (OR = 1.6 [0.8–3.3], p = 0.18) [22].

In contrast, *IL28B* polymorphism was associated with overall SVR in a cohort of Italian patients treated with variable-duration pegIFN and RBV for genotype 2/3 HCV [81], although the absolute effect size was attenuated compared to genotype 1 HCV (rs12979860, C/C: SVR = 82%% versus C/T = 75% versus T/T = 58%; $p < 0.01$). *IL28B* genotype was most relevant for predicting SVR in patients who failed to attain a week 4 RVR (39% of the overall cohort). In this non-RVR group the respective SVR rates were C/C = 78% versus C/T = 58% versus T/T = 32%; $p < 0.001$. An association between *IL28B* genotype, early virological kinetics, and RVR (but not necessarily SVR) was subsequently demonstrated by multiple replication studies [13,14,16,82–85]. This may reflect inadequate power to differentiate statistically significant differences where SVR rates are generally high.

Systematic differences between study populations—patient ethnicity, *IL28B* genotype frequency, ratio of genotype 2:3 HCV, as well as treatment regimen (short-duration, variable RBV dosing)—complicate comparisons of the multiple datasets. Larger prospective studies will therefore be necessary to define the role of *IL28B* genotyping in these patients. Most importantly, the relevance of *IL28B* genotype to non-RVR HCV2/3 requires validation. The data suggest that 24 weeks of treatment may be sufficient for C/C non-RVR patients, whereas T/T patients might benefit from extending to 48 weeks (note that the EASL guidelines recently recommended extending therapy to 48 weeks for non-RVR HCV-2/3 patients [79]). This warrants prospective evaluation. For now *IL28B* genotyping should remain a research tool in patients infected with genotype 2/3 HCV.

There are fewer data for other HCV genotypes. The standard treatment for genotype 4 HCV (HCV-4) is 48 weeks of pegIFN and RBV. Treatment response rates are similar to those observed for HCV-1. A number of small datasets suggest that in patients with HCV-4 infection, the predictive power of *IL28B* polymorphism is similar to that for genotype 1 HCV [22,86]. More detailed confirmatory studies of the relevance of *IL28B* genotype to HCV-4 infection, particularly in the Egyptian population, are required. The relevance to genotype 6 HCV is less clear. There is only one description of a small cohort of 24 patients of Chinese ancestry, in which 23 of 24 patients carried the good-response genotype for rs12979860 and all 23 attained an SVR. One patient carried a poor-response genotype and relapsed after 48 weeks of treatment [31].

HCV-HIV Coinfection

The association of *IL28B* genotype and HCV infection outcome does not seem to be affected by HIV coinfection. *IL28B* polymorphism is a strong predictor of pegIFN and RBV treatment response in coinfected patients with genotype HCV-1/4, with effect sizes that are similar to those in monoinfection [25,86–91]. *IL28B* genotyping may therefore be useful in planning the timing of antiviral therapy in genotype 1/4 HCV, with the caveat that there are as yet limited data on the use of DAAs in HIV coinfected patients and availability may be delayed. In genotype HCV-2/3 coinfection, the association between *IL28B* and IFN response is attenuated, as with HCV-2/3 mono-infection [86]. The association between *IL28B* polymorphism and spontaneous HCV clearance is not attenuated by HIV coinfection, where rates of clearance are not significantly different from those seen in HCV monoinfected individuals [8,22]. *IL28B* polymorphisms have been shown not to influence the natural history of HIV infection itself [92–95].

Post-Transplantation HCV

Recurrent HCV post–liver transplantation is universal and typically aggressive, leading to progressive fibrosis and risk of graft failure. Its eradication is associated with improved clinical outcomes [96–98]. Unfortunately, treatment with pegIFN and RBV is more complicated than in the nontransplant setting, with lower overall safety and efficacy [99]. DAA therapy has yet to be evaluated in the post-transplant setting, where drug–drug interactions are challenging [100], and is therefore likely to be some years away. For all of these reasons, it would be clinically useful to be able to more accurately identify those patients most likely to respond to IFN-based therapy.

A number of small retrospective studies have now evaluated the relationship between recipient and donor liver *IL28B* genotype and IFN treatment response for recurrent HCV post–liver transplantation [33,101–104]. The data indicate that both recipient and donor liver good-response *IL28B* genotypes are associated with an increased rate of SVR and that this benefit is additive. In a Japanese study, 41 patients were classified on the basis of recipient/donor *IL28B* genotype (good response (+) versus poor response (−)) [101]. SVR rates according to *IL28B* genotype pairs were: −/− 0% versus +/− 10% versus +/+ 56% ($p<0.01$). In a retrospective study of 65 consecutive patients from the Mayo clinic, SVR rates were: −/− 16% versus +/− 47% versus +/+ 86% ($p<0.01$) [33]. The data suggest that *IL28B* genotyping might be used to prioritize patients for treatment with pegIFN and RBV. Given the poor prognosis of post-transplant HCV, the data also suggest a theoretical benefit for preferential organ allocation according to *IL28B* genotype, notwithstanding the considerable logistic challenges this would present.

Acute HCV Infection

Good-response *IL28B* patients are more likely to clear HCV infection spontaneously [22–25]. Very recently, the same observation was made using the novel *IFNL4* polymorphism [45]. Good-response patients are also more likely to present with jaundice at the time of infection [23,24], which is consistent with the historical observation that jaundiced patients are more likely to clear HCV. Jaundice is not completely explained by *IL28B* genotype, however. In an analysis of 136 women from the German Anti-D cohort, an interaction was observed between *IL28B* genotype, jaundice, and spontaneous clearance [23]. In patients with the good-response *IL28B* genotype, jaundice was more common (33% versus 16%, $p=0.03$), but the rate of clearance was high even in the absence of jaundice (61% versus 56%, $p=0.98$). In contrast, in patients with the poor-response *IL28B* genotypes jaundice was less common, but it was predictive of clearance (43% versus 14%, $p=0.019$) [23]. Therefore, patients who carry the poor response *IL28B* variant, and who do not present with jaundice, are unlikely to spontaneously clear the infection.

The current recommendation for acute HCV infection management is treatment after a period of observation to allow for spontaneous clearance—typically 8–12 weeks from the onset of symptoms [79,105]. Treatment of acute HCV with pegIFN monotherapy is associated with cure rates that are considerably higher than for chronic hepatitis C. SVR rates >80–90% have been reported [79]. Not surprisingly given these response rates, no significant association between *IL28B* genotype and treatment outcome has been observed in this setting [24,25]. Patients carrying the poor-response *IL28B* variants appear to gain a particular advantage with treatment in the acute setting. The data on this have implications for management with pegIFN therapy. An observation period is very reasonable in patients carrying the good-response *IL28B* genotype, where the spontaneous clearance rate is >50% [8,22,23] and, in the event of persistent high SVR rates, is expected with combined pegIFN/RBV therapy. In contrast, patients carrying poor-response *IL28B* genotypes are unlikely to naturally clear infection, particularly if anicteric [23]. These patients might be considered for immediate therapy to maximize the likelihood of SVR.

37.5.2 Direct-Acting Antiviral Therapy for Genotype 1 HCV

37.5.2.1 *IL28B* Genotype and First-Generation Protease Inhibitors

The standard-of-care treatment for genotype 1 HCV is now the combination of an HCV-NS3 protease inhibitor (PI) and pegIFN/RBV. There are currently two PIs on the market, telaprevir (TVR) and boceprevir (BOC). Retrospective analyses of the association between *IL28B* genotype and treatment outcome in phase 3 TVR/BOC registration programs have shown that *IL28B* genotype remains relevant to treatment response in treatment-naïve patients [106,107]. However, the strength of the association is attenuated when compared to pegIFN and RBV response, because TVR/BOC significantly increase SVR rates in patients who carry the poor-response *IL28B* genotype [106,107]). The increment in SVR rate in good-response *IL28B* genotype patients was lower in these retrospective analyses (Table 37.1). In treatment-naïve patients, the SVR rate was no different in the BOC arms from that of controls among good-response *IL28B* patients (Table 37.1) [107]. A small increment in SVR in good-response *IL28B* patients due to TVR was observed (Table 37.1).

Importantly, the good-response *IL28B* genotype strongly predicted eligibility for short-duration (24–28 week) therapy in both studies (BOC: 89% versus 52% in CC versus non-CC patients, rs12979860 [108]; TVR: 78% versus 54% [109]). This shows that, for most good-response *IL28B* patients, TVR/BOC allows short-duration therapy. In poor-response *IL28B* patients, TVR/BOC regimens increase SVR rates and allow short-duration therapy in a subset. Even shorter therapy appears to be possible in good-response patients—it was recently shown that 12 weeks total duration of TVR-based triple therapy is associated with very high SVR rates [110].

IL28B genotyping is less useful for predicting the outcome of TVR/BOC therapy in patients who previously failed to respond to pegIFN/RBV dual therapy [106,107]. Here, the treatment history defines IFN responsiveness (relapse versus partial response versus nonresponse). *IL28B* genotyping no longer predicts for SVR once patients have been classified on this basis (Figure 37.3), but it may still have a role in real-world populations where previous IFN response is less well defined.

37.5.2.2 *IL28B* Genotype and Next-Generation DAA Response

The treatment of HCV is rapidly evolving. Data suggest that treatment-naïve patients who carry the good-response *IL28B* genotype are easier to cure with regimens that involve a single DAA combined with a pegIFN/RBV backbone, including next-wave NS3 protease inhibitors (e.g., simeprevir), NS5B inhibitors (e.g., sofosbuvir), and NS5A inhibitors (e.g., daclatasvir) [111–118]. The strength of association here is weaker than for pegIFN/RBV alone, however, and appears to negatively correlate with DAA antiviral potency. Indeed, as SVR rates reach very high levels, the clinical significance of small but statistically real differences in response according to *IL28B* genotype becomes less and less meaningful. The highest SVR rate observed to date for a triple therapy regimen in a phase 3 study was that reported for the 12-week regimen of sofosbuvir plus pegIFN/RBV [119]. The overall SVR rate was 89% in genotype 1 HCV patients, with SVR rates of 98% versus 87% according to *IL28B* genotype (rs12979860 CC versus non-CC, $p<0.05$) [119].

Quadruple therapy regimens with combination DAAs plus pegIFN/RBV are in phase 2 development. Early data show very high SVR rates in small cohorts [113], including null-responder cohorts, indicating that *IL28B* genotype may be less relevant in this setting. Thus, it is reasonable to expect little, if any, relationship between *IL28B* genotype and of IFN-free outcomes. One scenario in which both *IL28B* genotype and IFN remain relevant to future HCV therapy is if distinct treatment paradigms emerge for patients who are good versus poor IFN responders. Patients carrying the *IL28B* good-response genotype might be treated with dual DAA plus pegIFN/RBV for very short duration (e.g., 6–8 weeks) [120]. In contrast, poor-response *IL28B* genotype patients might require therapy of a longer duration or be prioritized for IFN-free therapy.

Interferon-free therapy for HCV-1 is likely to be approved in the next 3–5 years. Data suggest that the good-response *IL28B* genotype may represent an “easy-to-cure” characteristic for certain IFN-free regimens as well. The INFORM-1 study was the first to demonstrate that IFN-free therapy can have a potent antiviral effect. Patients were treated with a combination of mericitabine (a NS5B nucleotide inhibitor, NI) and danoprevir (a PI) for 14 days before follow-on PR therapy to 48 weeks. Analysis of the

TABLE 37.1 Retrospective Analyses of the Association Between *IL28B* Genotype and Treatment Response

Outcome		Treatment Regimen	*IL28B* Genotype (rs12979860)					
			C/C		C/T		T/T	
			No.	%	No.	%	No.	%
Boceprevir (BOC)								
Treatment-naïve [76] (n = 653/1048)								
SVR		BOC	44/55	80	82/115	71	26/44	59
		BOC RGT	63/77	82	67/103	65	23/42	55
		PR control	50/64	78	33/116	28	10/37	27
Week 8 response[a]		All BOC patients	118/132	89	158/304	—	—	52
Treatment-experienced [75] (n = 259/393)								
SVR		BOC	17/22	77%	48/66	73%	13/18	72
		BOC RGT	22/28	79	38/62	61	6/11	55
		PR control	6/13	46	5/29	17	5/10	50
Week 8 response[a]		All BOC patients	41/50	82	80/156	—	—	51
Telaprevir (TVR)								
Treatment-naïve[77] (n = 454/1088)								
SVR		T12	45/50	90	48/68	71	16/22	73
		T8	38/45	84	43/76	57	19/32	59
		PR control	35/55	64	20/80	25	6/26	23
eRVR[b]		T12	39/50	78	39/68	57	10/22	45
REALIZE [78](n = 527 / 662)								
SVR	Prior relapsers	All TVR patients	51/58	88	100/117	85	29/34	85
		PR control	4/12	33	6/30	20	3/10	30
	Prior partial responders	All TVR patients	5/8	63	33/57	58	10/14	71
		PR control	1/5	20	2/10	20	0/5	0
	Prior null responders	All TVR patients	4/10	40	27/92	29	10/32	31
		PR control	0/0	0	1/18	6	1/15	7

[a] *Patients were eligible for short-duration therapy if HCV RNA was undetectable at week 8 and week 24.* [b] *Patients were eligible for short-duration therapy if they achieved an extended rapid virological response (eRVR), defined by undetectable HCV RNA at weeks 4 and 12 of treatment.*

Note: The information here is from the phase 3 registration studies of boceprevir (BOC) and telaprevir (TVR) in both treatment-naïve [76,77] and treatment-experienced patients [75,78].

on-treatment viral kinetics in 15 patients during the two weeks of oral therapy revealed a significant difference in phase II viral kinetics according to *IL28B* genotype [121], suggesting that *IL28B* genotype influences the rate of clearance of infected hepatocytes during IFN-free therapy. This might be consistent with the association between *IL28B* genotype and spontaneous HCV clearance.

SOUND-C2 was the first large study of IFN-free therapy [122]. It evaluated the combination of faldaprevir (PI), deleobuvir (NNI) ± RBV (Figure 37.1). Interim results demonstrated a clear difference in SVR12 according to *IL28B* genotype (58–84% among patients with *IL28B* CC versus 33–64% of those with non-CC genotypes). The difference was more marked in patients infected with genotype 1a HCV versus genotype 1b. The low SVR rates observed in HCV-1a non-C/C patients resulted from a virological breakthrough in most patients, suggesting that *IL28B* genotype influences RAV emergence. It should be noted that this area is rapidly developing. Multiple IFN-free regimens that involve DAAs with high potency and a lower genetic barrier to resistance are now in phase 3 development. It is expected that response rates with "optimized" DAA combinations will be higher still and, again, will overcome most of the clinical impact of *IL28B* genotype.

A recent study on the combination of sofosbuvir plus daclatasvir reported SVR4 of 100% in 44 HCV-1 treatment-naïve patients with or without ribavirin [123]. Such optimized regimens are very expensive, however, and may not be necessary for all patients. *IL28B* genotype may continue to be useful in identifying patients for whom such regimens are most appropriate as first-line therapy.

In summary, it seems that it will be possible to overcome the effect of *IL28B* polymorphism in most patients with a combination of drugs with high potency and high genetic barriers to resistance (e.g., sofosbuvir + daclatasvir) ± pegIFN/RBV. Even so, it may be that *IL28B* genotype will be used to individualize treatment strategies such that patients who carry the good-response *IL28B* genotype may be eligible for shorter, simpler, or cheaper regimens and that, conversely, "poor-response" patients may require longer therapy with multiple DAAs. Such individualized treatment algorithms have yet to be explored in detail [110,120].

IL28B Genotype and Pharmaco-Economics

Antiviral therapy for HCV is expensive. The pricing schedule for telaprevir in North America in 2011 was $US49,200 for 12 weeks of treatment; for boceprevir, it was $US1100 per week (24–44 weeks). This comes on top of the cost for pegIFN and RBV. With finite healthcare budgets, it is important to ask whether prioritizing certain patients for these new agents is justified. *IL28B* genotyping is an obvious tool for use in this situation. Indeed, in the phase 3 registration study of the PI boceprevir, although triple therapy allowed shortening of total treatment duration in most patients, overall SVR rates did not increase relative to pegIFN and RBV among good-response *IL28B* patients [107]. One strategy would therefore be to treat all good-response *IL28B* genotypes with pegIFN and RBV alone, reserving DAAs for the minority of nonresponders [28,124,125]. Poor-response *IL28B* genotype patients might be prioritized for first-line access—a policy that has been adopted in a number of regions. In any case, formal economic modeling should be performed on a region-by-region basis, taking into consideration *IL28B* genotype, other baseline predictors of SVR such as fibrosis stage, and on-treatment milestones such as week 4 virological response.

37.5.3 Clinical Translation of *ITPA* Genotyping

ITPA genotype is strongly associated with ribavirin-induced anemia; predicted ITPase deficiency identifies approximately 25% of patients to be protected at week 4 [51]. Studies have confirmed protection throughout the 24–48-week course of pegIFN/RBV treatment [56,57]. Although patients with ITPase deficiency are less likely to require RBV dose reduction [50,51], *ITPA* variants are not associated with SVR [51,56,57]. Therefore, routine *ITPA* genotyping cannot at this time be recommended prior to pegIFN/RBV therapy. Particular patient groups exist in which *ITPA* genotype may be relevant to clinical decision-making. These include patients at high risk for anemia, such as thalassemic patients, or patients on renal replacement therapy. *ITPA* genotyping may also be relevant to patients who poorly tolerate anemia, such as those with coronary artery disease. Prospective evaluation will be important.

Both telaprevir- and boceprevir-based therapies are associated with more profound anemia compared to pegIFN and RBV alone [109,126]. The mechanism remains unclear, but is additive with RBV and appears to involve bone marrow suppression rather than hemolysis. *ITPA* genotype is associated with the risk of anemia during TVR therapy [127]. It is also associated with anemia, as well as the need for both RBV dose reduction and erythropoietin support, during BOC therapy [83]. *ITPA* genotype has not been shown to be causally associated with telaprevir/boceprevir anemia, however, and it appears to reflect the RBV effect alone. Effect size is small—it is moderate only for both regimens—and ITPA genotyping is not routinely recommended. Nevertheless, knowledge of *ITPA* genotype might be useful for clinicians when planning DAA therapy, especially in regions where access to erythropoietin (EPO) support is restricted or in subgroups at higher risk for anemia (e.g., older patients).

37.5.4 Genomics and Drug Development

Current strategies for the clinical development of DAAs involve combination with pegIFN and RBV to restrict the development of antiviral resistance. Given the population frequency of the good-response *IL28B* variant, and its association with rapid viral decline, it is possible for small early-phase efficacy trials to be confounded by imbalance for *IL28B* genotype across treatment arms. For this reason, treatment arms should be stratified by *IL28B* genotype [128]. This approach has largely been adopted by the pharmaceutical industry [129].

Beyond stratification, pharmacogenetic analyses have the potential to redefine drug development. Analysis for genetic predictors of treatment efficacy as well as treatment toxicity will be important as a means of enhancing efficacy and minimizing toxicity. Future drug development may be targeted accordingly—for example, where a genetic marker for toxicity is identified in early-phase studies, drug candidates may advance into later-phase studies that, rather than halt development entirely, enroll only individuals with low-risk variants. Telaprevir rash is an obvious case where identification of a genetic susceptibility marker might have proved useful during the development phase; however, telaprevir still warrants investigation in the post-regulatory period. Most studies of novel anti-HCV drugs now include an optional pharmacogenomics substudy. This approach has potential application in all fields of therapeutic development.

37.6 CONCLUSION

The discoveries of the associations of *IL28B* and *ITPA* polymorphisms with HCV outcomes are two of the most definitive successes of the GWAS era [130,131]—personalized medicine is now a reality for chronic HCV patients. Nevertheless, such important discoveries should not obscure the intrinsic limitations of GWASs: for one, the stringent statistical significance threshold that must be used to correct for multiple testing severely limits the power to uncover variants that have low to moderate effect size. For another, by using tagging SNPs, GWASs only interrogate common genetic variants and are for the most part uninformative about the bulk of rarer DNA changes that can have a dramatic impact at the individual level, as seen in many Mendelian disorders. It must also be remembered that GWASs identify loci, not genes, and the identification of pathogenic changes (causal variants/mechanism) on an associated haplotype can be complicated. In fact, it is not unreasonable to argue that GWASs thus far represent only a temporary stop on the road toward more comprehensive HCV host genomic studies: next-generation sequencing technology is now widely available and permits cost-effective sequencing of whole human exomes, genomes, or transcriptomes.

Detailed characterization of individual DNA/RNA profiles, in combination with systems approaches for statistical analysis, has the potential to shed new light on the complex interplay between HCV and the host cell, and to provide novel therapeutic targets. Computational issues need to be addressed, and novel statistical methods will be required to make the best use of the unprecedented wealth of genomic information that is being generated. Still, the promise is, finally, a fully comprehensive representation of all human genetic variation. These are exciting times, indeed, for human genomic research, including the host genetics of HCV disease.

ACKNOWLEDGMENTS

A.J.T. has received funding support from the National Health and Medical Research Council of Australia and the University of Melbourne. P.J.C has received funding from the National Health and Medical Research Council of Australia.

DISCLOSURES

A.J.T is a coinventor of patents related to the *IL28B* and *ITPA* discoveries. He has also served on advisory committees or review panels (Merck, Janssen, Roche, Gilead, Glaxo-Smith-Kline, Abbvie, Bristol-Myers Squibb, Springbank), has held speaking or teaching roles (Merck, Roche, Bristol-Myers Squibb, Janssen, Gilead), and has received grant/research support (Merck, Roche, Gilead).

P.J.C. has held speaking or teaching roles (Janssen).

A.A. has served on advisory boards or review panels (Roche, Janssen), has held speaking or teaching roles (Roche, Janssen, Merck, Bayer), and has received grant/research support (Roche, Gilead).

ABBREVIATIONS

IL	interleukin
ITP	inosine triphosphate
ITPA	inosine triphosphatase
HCV	hepatitis C virus
GWAS	genome-wide association study
pegIFN	pegylated interferon-α, RBV, ribavirin
SNP	single-nucleotide polymorphism
LD	linkage disequilibrium
OR	odds ratio
HIV	human immunodeficiency virus
ISG	interferon-stimulated gene
Hb	hemoglobin
RVR	rapid virological response
CEVR	complete early virological response
SVR	sustained virological response
PPV	positive predictive value
DAA	direct-acting antiviral therapy
TVR	telaprevir
BOC	boceprevir
PI	protease inhibitor
NI	nucleotide inhibitor
NNI	nonnucleos(t)ide inhibitor
IFNL	*interferon-lambda*

REFERENCES

[1] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.

[2] International HapMap Consortium. A haplotype map of the human genome. Nature 2005;437:1299–320.

[3] McHutchison JG, Lawitz EJ, Shiffman ML, Muir AJ, Galler GW, McCone J, et al. Peginterferon alfa-2b or alfa-2a with ribavirin for treatment of hepatitis C infection. N Engl J Med 2009;361:580–93.

[4] Muir AJ, Bornstein JD, Killenberg PG. Peginterferon alfa-2b and ribavirin for the treatment of chronic hepatitis C in blacks and non-Hispanic whites. N Engl J Med 2004;350:2265–71.

[5] Thompson AJ, Muir AJ, Sulkowski MS, Ge D, Fellay J, Shianna KV, et al. Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. Gastroenterology 2010;139(120–9):120–9 [e18].

[6] Bochud PY, Bibert S, Negro F, Haagmans B, Soulier A, Ferrari C, et al. IL28B polymorphisms predict reduction of HCV RNA from the first day of therapy in chronic hepatitis C. J Hepatol 2011;55:980–8.

[7] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461:399–401.

[8] Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. Nature 2009;461:798–801.

[9] Liu CH, Liu CJ, Lin CL, Liang CC, Hsu SJ, Yang SS, et al. Pegylated interferon-alpha-2a plus ribavirin for treatment-naive Asian patients with hepatitis C virus genotype 1 infection: a multicenter, randomized controlled trial. Clin Infect Dis 2008;47:1260–9.

[10] Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 2009;41(10):1100–4.

[11] Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 2009;41:1105–9.

[12] McCarthy JJ, Li JH, Thompson A, Suchindran S, Lao XQ, Patel K, et al. Replicated association between an IL28B gene variant and a sustained response to pegylated interferon and ribavirin. Gastroenterology 2010;138:2307–14.

[13] Hezode C, Hirschfield GM, Ghesquiere W, Sievert W, Rodriguez-Torres M, Shafran SD, et al. BMS-790052, a NS5A replication complex inhibitor, combined with peginterferon alfa-2a and ribavirin ion treatment-naive HCV-genotype 1 or 4 patients: phase 2B AI444010 study interim week 12 results. Hepatology 2011;54:474A [abstract 227].

[14] Zeuzem S, Soriano V, Asselah T, Bronowicki J-P, Lohse AW, Mullhaupt B, et al. SVR4 and SVR12 with an interferon-free regimen of BI 201335 and BI 207127, +/- ribavirin, in treatment-naïve patients with chronic genotype-1 HCV infection: interim results of SOUND-C2. J Hepatol 2012:56 [abstract 101].

[15] Chayama K, Hayes CN, Abe H, Miki D, Ochi H, Karino Y, et al. IL28B but not ITPA polymorphism is predictive of response to pegylated interferon, ribavirin, and telaprevir triple therapy in patients with genotype 1 hepatitis C. J Infect Dis 2011;204:84–93.

[16] Hardy J, Singleton A. Genomewide association studies and human disease. N Engl J Med 2009;360:1759–68.

[17] Lin CY, Chen JY, Lin TN, Jeng WJ, Huang CH, Huang CW, et al. IL28B SNP rs12979860 is a critical predictor for on-treatment and sustained virologic response in patients with hepatitis C virus genotype-1 infection. PLoS One 2011;6:e18322.

[18] Lagging M, Askarieh G, Negro F, Bibert S, Söderholm J, Westin J, et al. For the D-HCVSG. Response prediction in chronic hepatitis C by assessment of IP-10 and IL28B related Single nucleotide polymorphisms. PLoS One 2011;6:e17232.

[19] Ochi H, Hayes CN, Abe H, Hayashida Y, Uchiyama TK, Kamatani N, et al. Towards the establishment of a prediction system for the personalized treatment of chronic hepatitis C. J Infect Dis 2012;205:204–10.

[20] Hayes CN, Kobayashi M, Akuta N, Suzuki F, Kumada H, Abe H, et al. HCV substitutions and IL28B polymorphisms on outcome of peg-interferon plus ribavirin combination therapy. Gut 2011;60:261–7.

[21] Hiraga N, Abe H, Imamura M, Tsuge M, Takahashi S, Hayes CN, et al. Impact of viral amino acid substitutions and host IL28B polymorphism on replication and susceptibility to interferon of hepatitis C virus. Hepatology 2011;54(3):764–71.

[22] Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. Gastroenterology 2010;138:1338–45. 1345 [e1–7].

[23] Tillmann HL, Thompson AJ, Patel K, Wiese M, Tenckhoff H, Nischalke HD, et al. A polymorphism near IL28B is associated with spontaneous clearance of acute hepatitis C virus and jaundice. Gastroenterology 2010;139:1586–92. [1592 e1]

[24] Grebely J, Petoumenos K, Hellard M, Matthews GV, Suppiah V, Applegate T, et al. Potential role for interleukin-28B genotype in treatment decision-making in recent hepatitis C virus infection. Hepatology 2010;52:1216–24.

[25] Nattermann J, Vogel M, Nischalke HD, Danta M, Mauss S, Stellbrink HJ, et al. Genetic variation in IL28B and treatment-induced clearance of hepatitis C virus in HIV-positive patients with acute and chronic hepatitis C. J Infect Dis 2011;203:595–601.

[26] Thompson AJ, Clark PJ, Zhu M, Zhu Q, Ge D, Sulkowski MS, et al. Genome wide-association study GWAS identifies IL28B polymorphism to be associated with baseline ALT and hepatic necro-inflammatory activity in chronic hepatitis C patients enrolled in the IDEAL study. Hepatology 2010;52:1893A.

[27] Thompson AJ, Clark PJ, Fellay J, Muir AJ, Tillmann HL, Patel K, et al. IL28B genotype is not associated with advanced hepatic fibrosis in chronic hepatitis C patients enrolled in the IDEAL study. Hepatology 2010;52:229A.

[28] Gellad ZF, Naggie S, Reed SD, Clark PJ, Thompson AJ, Schulman KA, et al. The cost-effectiveness of a telaprevir-inclusive regimen as initial therapy for genotype 1 hepatitis C infection in individuals with the C/C IL28B polymorphism. Hepatology 2011;54 (Suppl. 4):417A [118].

[29] The International HapMap Consortium. A haplotype map of the human genome. Nature 2005;437:1299–320.

[30] Abe H, Ochi H, Maekawa T, Hayes CN, Tsuge M, Miki D, et al. Common variation of IL28 affects gamma-GTP levels and inflammation of the liver in chronically infected hepatitis C virus patients. J Hepatol 2010;53:439–43.

[31] Seto WK, Tanaka Y, Liu K, Lai CL, Yuen MF. The Effects of IL-28B and ITPA polymorphisms on treatment of hepatitis C virus genotype 6. Am J Gastroenterol 2011;106:1007–8.

[32] Zagordi O, Geyrhofer L, Roth V, Beerenwinkel N. Deep sequencing of a genetically heterogeneous sample: local haplotype reconstruction and read error correction. J Comput Biol 2010;17:417–28.

[33] Charlton MR, Thompson A, Veldt BJ, Watt K, Tillmann H, Poterucha JJ, et al. Interleukin-28B polymorphisms are associated with histological recurrence and treatment response following liver transplantation in patients with hepatitis C virus infection. Hepatology 2011;53:317–24.

[34] Asselah T. A sprint to increase response to HCV treatment: expectancies but caution. J Hepatol 2011;55:1154–8.

[35] Tillmann HL, Patel K, Muir AJ, Guy CD, Li JH, Lao XQ, et al. Beneficial IL28B genotype associated with lower frequency of hepatic steatosis in patients with chronic hepatitis C. J Hepatol 2011;55(6):1195–200.

[36] Clark PJ, Thompson AJ, Zhu M, Zhu Q, Ge D, Sulkowski MS, et al. IL28B genetic polymorphism has genome-wide significant association with serum low-density lipoprotein levels and is strongly correlated with hepatic steatosis in patients with genotype 1 chronic hepatitis C. Hepatology 2010;52:232A.

[37] Aghemo A, Rumi MG, De Nicola S, Colombo M. Twelve-week posttreatment follow-up predicts a sustained virological response to pegylated interferon and ribavirin therapy. Hepatology 2010;52:1170–1.

[38] Poynard T, McHutchison J, Manns M, Myers RP, Albrecht J. Biochemical surrogate markers of liver fibrosis and activity in a randomized trial of peginterferon alfa-2b and ribavirin. Hepatology 2003;38:481–92.

[39] http://investors.gilead.com/phoenix.zhtml?c=69964&p=irol-newsArticle&ID=1662331&highlight=2012.

[40] Urban TJ, Thompson AJ, Bradrick SS, Fellay J, Schuppan D, Cronin KD, et al. IL28B genotype is associated with differential expression of intrahepatic interferon-stimulated genes in patients with chronic hepatitis C. Hepatology 2010;52:1888–96.

[41] Honda M, Sakai A, Yamashita T, Nakamoto Y, Mizukoshi E, Sakai Y, et al. Hepatic ISG expression is associated with genetic variation in IL28B and the outcome of IFN therapy for chronic hepatitis C. Gastroenterology 2010;139:499–509.

[42] Dill MT, Duong FH, Vogt JE, Bibert S, Bochud PY, Terracciano L, et al. Interferon-induced gene expression is a stronger predictor of treatment response than IL28B genotype in patients with hepatitis C. Gastroenterology 2011;140:1021–31 [e10].

[43] Naggie S, Osinusi A, Katsounas A, Lempicki R, Herrmann E, Thompson AJ, et al. Dysregulation of innate immunity in hepatitis C virus genotype 1 IL28B-unfavorable genotype patients: impaired viral kinetics and therapeutic response. Hepatology 2012;56:444–54.

[44] Prokunina-Olsson L, Muchmore B, Tang W, Pfeiffer RM, Park H, Dickensheets H, et al. A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. Nat Genet 2013;45:164–71.

[45] Aka P, Kuniholm MH, Pfeiffer RM, Wang AS, Tang W, Chen S, et al. Association of the IFNL4-deltaG allele with impaired spontaneous clearance of hepatitis C virus. J Infect Dis 2014;209(3):350–4.

[46] Smith KR, Suppiah V, O'Connor K, Berg T, Weltman M, Abate ML, et al. Identification of improved IL28B SNPs and haplotypes for prediction of drug response in treatment of hepatitis C using massively parallel sequencing in a cross-sectional European cohort. Genome Med 2011;3:57.

[47] Fischer J, Bohm S, Scholz M, Muller T, Witt H, George J, et al. Combined effects of different interleukin-28B gene variants on the outcome of dual combination therapy in chronic hepatitis C virus type 1 infection. Hepatology 2012;55:1700–10.

[48] Kotenko SV, Gallagher G, Baurin VV, Lewis-Antes A, Shen M, Shah NK, et al. IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. Nat Immunol 2003;4:69–77.

[49] Muir AJ, Shiffman ML, Zaman A, Yoffe B, de la Torre A, Flamm S, et al. Phase 1b study of pegylated interferon lambda 1 with or without ribavirin in patients with chronic genotype 1 hepatitis C virus infection. Hepatology 2010;52(3):822–32.

[50] Sarasin-Filipowicz M, Oakeley EJ, Duong FH, Christen V, Terracciano L, Filipowicz W, Heim MH. Interferon signaling and treatment outcome in chronic hepatitis C. Proc Natl Acad Sci USA 2008;105:7034–9.

[51] Fellay J, Thompson AJ, Ge D, Gumbs CE, Urban TJ, Shianna KV, et al. ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. Nature 2010;464:405–8.

[52] Sumi S, Marinaki AM, Arenas M, Fairbanks L, Shobowale-Bakre M, Rees DC, et al. Genetic basis of inosine triphosphate pyrophosphohydrolase deficiency. Hum Genet 2002;111:360–7.

[53] Shipkova M, Lorenz K, Oellerich M, Wieland E, von Ahsen N. Measurement of erythrocyte inosine triphosphate pyrophosphohydrolase (ITPA) activity by HPLC and correlation of ITPA genotype-phenotype in a Caucasian population. Clin Chem 2006;52:240–7.

[54] Atanasova S, Shipkova M, Svinarov D, Mladenova A, Genova M, Wieland E, et al. Analysis of ITPA phenotype-genotype correlation in the Bulgarian population revealed a novel gene variant in exon 6. Ther Drug Monit 2007;29:6–10.

[55] Maeda T, Sumi S, Ueta A, Ohkubo Y, Ito T, Marinaki AM, et al. Genetic basis of inosine triphosphate pyrophosphohydrolase deficiency in the Japanese population. Mol Genet Metab 2005;85:271–9.

[56] Thompson AJ, Fellay J, Patel K, Tillmann HL, Naggie S, Ge D, et al. Variants in the ITPA gene protect against ribavirin-induced hemolytic anemia and decrease the need for ribavirin dose reduction. Gastroenterology 2010;139:1181–9.

[57] Thompson AJ, Santoro R, Piazzolla V, Clark PJ, Naggie S, Tillmann HL, et al. Inosine triphosphatase genetic variants are protective against anemia during antiviral therapy for HCV2/3 but do not decrease dose reductions of RBV or increase SVR. Hepatology 2011;53:389–95.

[58] Ochi H, Maekawa T, Abe H, Hayashida Y, Nakano R, Kubo M, et al. ITPA polymorphism affects ribavirin-induced anemia and outcomes of therapy–a genome-wide study of Japanese HCV virus patients. Gastroenterology 2010;139:1190–7.

[59] Tanaka Y, Kurosaki M, Nishida N, Sugiyama M, Matsuura K, Sakamoto N, et al. Genome-wide association study identified ITPA/DDRGK1 variants reflecting thrombocytopenia in pegylated interferon and ribavirin therapy for chronic hepatitis C. Hum Mol Genet 2011;20:3507–16.

[60] Hitomi Y, Cirulli ET, Fellay J, McHutchison JG, Thompson AJ, Gumbs CE, et al. Inosine triphosphate protects against ribavirin-induced adenosine triphosphate loss by restoring adenylosuccinate synthase function. Gastroenterology 2011;140(4):1314–21.

[61] Thompson AJ, Clark PJ, Singh A, Ge D, Fellay J, Zhu M, et al. Genome-wide association study of interferon-related cytopenia in chronic hepatitis C patients. J Hepatol 2012;56(2):313–9.

[62] Poynard T, Afdhal NH. Perspectives on fibrosis progression in hepatitis C: an a la carte approach to risk factors and staging of fibrosis. Antivir Ther 2010;15:281–91.

[63] Okamoto K, Mimura K, Murawaki Y, Yuasa I. Association of functional gene polymorphisms of matrix metalloproteinase (MMP)-1, MMP-3 and MMP-9 with the progression of chronic liver disease. J Gastroenterol Hepatol 2005;20:1102–8.

[64] Muhlbauer M, Bosserhoff AK, Hartmann A, Thasler WE, Weiss TS, Herfarth H, et al. A novel MCP-1 gene polymorphism is associated with hepatic MCP-1 expression and severity of HCV-related liver disease. Gastroenterology 2003;125:1085–93.

[65] Baur K, Mertens JC, Schmitt J, Iwata R, Stieger B, Eloranta JJ, et al. Combined effect of 25-OH vitamin D plasma levels and genetic vitamin D receptor (NR 1I1) variants on fibrosis progression rate in HCV patients. Liver Int 2012;32:635–43.

[66] Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461–5.

[67] Stickel F, Buch S, Lau K, Meyer zu Schwabedissen H, Berg T, Ridinger M, et al. Genetic variation in the PNPLA3 gene is associated with alcoholic liver injury in caucasians. Hepatology 2011;53:86–95.

[68] Trepo E, Pradat P, Potthoff A, Momozawa Y, Quertinmont E, Gustot T, et al. Impact of patatin-like phospholipase-3 (rs738409 C>G) polymorphism on fibrosis progression and steatosis in chronic hepatitis C. Hepatology 2011;54:60–9.

[69] Valenti L, Rumi M, Galmozzi E, Aghemo A, Del Menico B, De Nicola S, et al. Patatin-like phospholipase domain-containing 3 I148M polymorphism, steatosis, and liver damage in chronic hepatitis C. Hepatology 2011;53:791–9.

[70] Huang H, Shiffman ML, Friedman S, Venkatesh R, Bzowej N, Abar OT, et al. A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C. Hepatology 2007;46:297–306.

[71] Marcolongo M, Young B, Dal Pero F, Fattovich G, Peraro L, Guido M, et al. A seven-gene signature (cirrhosis risk score) predicts liver fibrosis progression in patients with initially mild chronic hepatitis C. Hepatology 2009;50:1038–44.

[72] Trepo E, Potthoff A, Pradat P, Bakshi R, Young B, Lagier R, et al. Role of a cirrhosis risk score for the early prediction of fibrosis progression in hepatitis C patients with minimal liver disease. J Hepatol 2011;55:38–44.

[73] do ON, Eurich D, Schmitz P, Schmeding M, Heidenhain C, Bahra M, et al. A 7-gene signature of the recipient predicts the progression of fibrosis after liver transplantation for hepatitis C virus infection. Liver Transpl 2012;18:298–304.

[74] Patin E, Kutalik Z, Guergnon J, Bibert S, Nalpas B, Jouanguy E, et al. Genome-wide association study identifies variants associated with progression of liver fibrosis from HCV infection. Gastroenterology 2012;143:1244–52 [e1–12].

[75] Bacon BR, Gordon SC, Lawitz E, Marcellin P, Vierling JM, Zeuzem S, et al. Boceprevir for previously treated chronic HCV genotype 1 infection. N Engl J Med 2011;364:1207–17.

[76] Poordad F, McCone Jr J, Bacon BR, Bruno S, Manns MP, Sulkowski MS, et al. Boceprevir for untreated chronic HCV genotype 1 infection. N Engl J Med 2011;364:1195–206.

[77] Jacobson IM, McHutchison JG, Dusheiko G, Di Bisceglie AM, Reddy KR, Bzowej NH, et al. Telaprevir for previously untreated chronic hepatitis C virus infection. N Engl J Med 2011;364:2405–16.

[78] Zeuzem S, Andreone P, Pol S, Lawitz E, Diago M, Roberts S, et al. Telaprevir for retreatment of HCV infection. N Engl J Med 2011;364:2417–28.

[79] Craxi A, Pawlotsky JM, Wedemeyer H, Bjoro K, Flisiak R, Forns X, et al. EASL Clinical Practice Guidelines: Management of hepatitis C virus infection. J Hepatol 2011;55:254–64.

[80] Liu CH, Liang CC, Liu CJ, Tseng TC, Lin CL, Yang SS, et al. Interleukin 28B genetic polymorphisms and viral factors help identify HCV genotype-1 patients who benefit from 24-week pegylated interferon plus ribavirin therapy. Antivir Ther 2012;17:477–84.

[81] Mangia A, Thompson AJ, Santoro R, Piazzolla V, Tillmann HL, Patel K, et al. An IL28B polymorphism determines treatment response of hepatitis C virus genotype 2 or 3 patients who do not achieve a rapid virologic response. Gastroenterology 2010;139:821–7 [827 e1].

[82] Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747–53.

[83] Sulkowski MS, Reddy KR, Pedicone LD, Shen J, Burroughs M, Brass CA, et al. ITPA deficiency is associated with lower rates of anemia and EPO use in patients treated with boceprevir (BOC) plus peginterferon/ribavirin (PR). Hepatology 2011;54:798A [934].

[84] Lawitz E, Lalezari JP, Hassanein T, Kowdley K, Poordad FF, Sheikh AM, et al. Once-daily PSI-7977 Plus PEG/RBV in treatment-naive patients with HCV Gt1: robust end-of-treatment response rates are sustained post-treatment. Hepatology 2011;54:472A [225].

[85] Liu C, Liang CC, Liu CJ, Tseng TC, Lin CL, Yang SS, et al. Interleukin 28B genetic polymorphisms and viral factors help identify hepatitis C virus genotype 1 patients who benefit from 24 week of Peginterferon plus Ribavirin therapy. Hepatology 2011;54:414A [564].

[86] Rallon NI, Naggie S, Benito JM, Medrano J, Restrepo C, Goldstein D, et al. Association of a single nucleotide polymorphism near the interleukin-28B gene with response to hepatitis C therapy in HIV/hepatitis C virus-coinfected patients. Aids 2010;24:F23–9.

[87] Pineda JA, Caruz A, Rivero A, Neukam K, Salas I, Camacho A, et al. Prediction of response to pegylated interferon plus ribavirin by IL28B gene variation in patients coinfected with HIV and hepatitis C virus. Clin Infect Dis 2010;51:788–95.

[88] Aparicio E, Parera M, Franco S, Perez-Alvarez N, Tural C, Clotet B, et al. IL28B SNP rs8099917 is strongly associated with pegylated interferon-alpha and ribavirin therapy treatment failure in HCV/HIV-1 coinfected patients. PLoS One 2010;5:e13771.

[89] Labarga P, Barreiro P, Mira JA, Vispo E, Rallon N, Neukam K, et al. Impact of IL28B polymorphisms on response to peginterferon and ribavirin in HIV-hepatitis C virus-coinfected patients with prior nonresponse or relapse. Aids 2011;25:1131–3.

[90] Rallon NI, Soriano V, Naggie S, Restrepo C, Goldstein D, Vispo E, et al. IL28B gene polymorphisms and viral kinetics in HIV/hepatitis C virus-coinfected patients treated with pegylated interferon and ribavirin. Aids 2011;25:1025–33.

[91] de Araujo ES, Dahari H, Cotler SJ, Layden TJ, Neumann AU, Melo CE, et al. Pharmacodynamics of PEG-IFN-[alpha]-2a and HCV response as a function of IL28B polymorphism in HIV/HCV-coinfected patients. J Acquir Immune Defic Syndr 2011;56:95–9.

[92] Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A whole-genome association study of major determinants for host control of HIV-1. Science 2007;317:944–7.

[93] Martin MP, Qi Y, Goedert JJ, Hussain SK, Kirk GD, Hoots WK, et al. IL28B polymorphism does not determine outcomes of hepatitis B virus or HIV infection. J Infect Dis 2010;202:1749–53.

[94] Rallon NI, Restrepo C, Naggie S, Lopez M, Del Romero J, Goldstein D, et al. Interleukin-28B gene polymorphisms do not influence the susceptibility to HIV-infection or CD4 cell decline. Aids 2011;25:269–71.

[95] Salgado M, Kirk GD, Cox A, Rutebemberwa A, Higgins Y, Astemborski J, et al. Protective interleukin-28B genotype affects hepatitis C virus clearance, but does not contribute to HIV-1 control in a cohort of African-American elite controllers/suppressors. Aids 2011;25:385–7.

[96] Veldt BJ, Poterucha JJ, Watt KD, Wiesner RH, Hay JE, Kremers WK, et al. Impact of pegylated interferon and ribavirin treatment on graft survival in liver transplant patients with recurrent hepatitis C infection. Am J Transplant 2008;8:2426–33.

[97] Zeuzem S, Poordad F, Gordon SC, Jacobson IM, Ralston R, Sniukiene V, et al. Boceprevir resistance-associated variants (RAVs) are observed more frequently in HCV (GT1) Infected patients with poor response to peginterferon alfa-2B/ribavirin. J Hepatol 2011;54:S4–S5.

[98] Rauch A, James I, Pfafferott K, Nolan D, Klenerman P, Cheng W, et al. Divergent adaptation of hepatitis C virus genotypes 1 and 3 to human leukocyte antigen-restricted immune pressure. Hepatology 2009;50:1017–29.

[99] Gaudieri S, Rauch A, Pfafferott K, Barnes E, Cheng W, McCaughan G, et al. Hepatitis C virus drug resistance and immune-driven adaptations: relevance to new antiviral therapy. Hepatology 2009;49:1069–82.

[100] Powdrill MH, Tchesnokov EP, Kozak RA, Russell RS, Martin R, Svarovskaia ES, et al. Contribution of a mutational bias in hepatitis C virus replication to the genetic barrier in the development of drug resistance. Proc Natl Acad Sci USA 2011;108:20509–13.

[101] Fukuhara T, Taketomi A, Motomura T, Okano S, Ninomiya A, Abe T, et al. Variants in IL28B in liver recipients and donors correlate with response to peg-interferon and ribavirin therapy for recurrent hepatitis C. Gastroenterology 2010;139:1577–85 [1585 e1–3].

[102] Lange CM, Moradpour D, Doehring A, Lehr HA, Müllhaupt B, Bibert S, et al. Impact of donor and recipient IL28B rs12979860 genotypes on hepatitis C virus liver graft reinfection. J Hepatol 2010;55:322–7.

[103] Eurich D, Boas-Knoop S, Ruehl M, Schulz M, Carrillo ED, Berg T, et al. Relationship between the interleukin-28b gene polymorphism and the histological severity of hepatitis C virus-induced graft inflammation and the response to antiviral therapy after liver transplantation. Liver transpl 2011;17:289–98.

[104] Coto-Llerena M, Perez-Del-Pulgar S, Crespo G, Carrion JA, Martinez SM, Sanchez-Tapias JM, et al. Donor and recipient IL28B polymorphisms in HCV-infected patients undergoing antiviral therapy before and after liver transplantation. Am J Transplant 2011;11:1051–7.

[105] Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 2009;49:1335–74.

[106] Pol S, Aerssens J, Zeuzem S, Andreone P, Lawitz EJ, Roberts S, et al. Limited impact of IL28B genotype on response rates in telaprevir-treated patients with prior treatment failure. J Hepatol 2013;58:883–9.

[107] Poordad F, Bronowicki JP, Gordon SC, Zeuzem S, Jacobson IM, Sulkowski MS, et al. Factors that predict response of patients with hepatitis C virus infection to boceprevir. Gastroenterology 2012;143(3):608–18.

[108] Poordad F, Bronowicki JP, Gordon S, Zeuzem S, Jacobson I, Sulkowski M, et al. IL28B polymorphism predicts virologic response in patients with hepatitis C genotype 1 treated with boceprevir (BOC) combination therapy. J Hepatol 2011;54:S6.

[109] Jacobson IM, Catlett I, Marcellin P, Bzowej NH, Muir AJ, Adda N, et al. Telaprevir substantially improved SVR rates across all IL28B genotypes in the ADVANCE trial. J Hepatol 2011:54 [abstract 1369].

[110] Nelson DR, Poordad F, Feld JJ, Fried MW, Jacobson IM, Pockros PJ, et al. High SVR rates (SVR4) for 12-week total telaprevir combination therapy in IL28B CC treatment-naives and prior relapsers with G1 chronic hepatitis C: CONCISE interim analysis. 48th annual meeting of the European association for the study of the liver; 2013 [abstract 881].

[111] Aerssens J, Fanning G, Scholliers A, Lenz O, Peeters M, De Srnedt G, et al. Impact of IL28B genotype and pretreatment serum IP-10 in treatment-naive genotype-1 HCV patients treated with TMC435 in combination with peginterferona-2A and ribavirin in pillar study. J Hepatol 2011;54:S5–S6 [A11].

[112] Zeuzem S, Foster GR, Fried MW, Hezode C, Hirschfield GM, Nikitin I, et al. The ASPIRE trial: TMC435 in treatment-experienced patients with genotype-1 HCV infection who have failed previous pegIFN/RBV treatment. J Hepatol 2011;54:A1376.

[113] Lok AS, Gardiner DF, Lawitz E, Martorell C, Everson G, Ghalib R, et al. Combination therapy with BMS-790052 and BMS-650032 alone or with pegIFN/RBV results in undetectable HCV RNA through 12 weeks of therapy in HCV genotype 1 null responders. J Hepatol 2011:54 [LB-8].

[114] Pol S, Ghalib VK, Rustgi C, Martorell GT, Everson HA, Tatum C, et al. First report of SVR12 for a NS5A replication complex inhibitor BMS-790052 in combination with PEG-IFNa-2A and RBV: phase 2A trial in treatment-naive HCV-genotype-1 subjects. J Hepatol 2011;54:S544–5 [A1373].

[115] Pockros P, Jensen D, Tsai N, Taylor RM, Ramji A, Cooper C, et al. First SVR data with the nucleoside analogue polymerase inhibitor mericitabine (RG7128) combined with peginterferon/ribavirin in treatment-naive HCV G1/4 patients: Interim analysis from the JUMP-C trial. J Hepatol 2011;54:A1359.

[116] Nelson DR, Lalezari J, Lawitz E, Hassanein T, Kowdley K, Poordad F, et al. Once daily PSI-7977 plus PEG-IFN/RBV in HCV GT1: 98% rapid virologic response, complete early virologic response: the proton study. J Hepatol 2011;54:S544 [A1372].

[117] Muir AJ, Lawitz E, Rodriguez-Torres M, Rustgi V, Hassanein T, Appleman JR, et al. IL28B polymorphism and kinetics of antiviral activity for ANA598 in combination with pegylated interferon alpha-2a plus ribavirin in treatment-naive genotype-1 chronic HCV patients. Hepatology 2010;52(S1) [abstract 1852].

[118] Flisiak R, Pawlotsky JM, Crabbé R, Calistru PI, Kryczka W, Haussinger D, et al. Once daily alisporivir (DEB025) plus pegIFNalfa2a/ribavirin results in superior sustained virologic response (SVR24) in chronic hepatitis C genotype 1 treatment naive patients. J Hepatol 2011;54:A4.

[119] Lawitz E, Mangia A, Wyles D, Rodriguez-Torres M, Hassanein T, Gordon SC, et al. Sofosbuvir for previously untreated chronic hepatitis C infection. N Engl J Med 2013;368:1878–87.

[120] Thompson A, Han S, Shiffman M, Rossaro L, Ghalib R, Beavers K, et al. GS-5885 + GS-9451 + Peginterferon and ribavrin (pr) for six or 12 weeks achieves a high SVR12 In treatment naïve genotype 1 IL28B CC PATIENTS. 48th annual meeting of the European association for the study of the liver; 2013 [abstract 64].

[121] Chu TW, Kulkarni R, Gane EJ, Roberts SK, Stedman C, Angus PW, et al. Effect of IL28B genotype on early viral kinetics during interferon-free treatment of patients with chronic hepatitis C. Gastroenterology 2012;142:790–5.

[122] Zeuzem S, Soriano V, Asselah T, Bronowicki JP, Lohse AW, Müllhaupt B, et al. Faldaprevir and deleobuvir for HCV genotype 1 infection. N Engl J Med 2013;369:630–9.

[123] Sulkowski MS, Gardiner DF, Rodriguez-Torres M, Reddy KR, Hassanein T, Jacobson I, et al. High rate of sustained virologic response with the all-oral combination of daclatasvir (NS5A inhibitor) plus sofosbuvir (nucleotide NS5B inhibitor), with or without ribavirin, in treatment-naive patients chronically infected with HCV genotype 1, 2, or 3. Program and abstracts of the 63rd annual meeting of the American association for the study of liver diseases; 2012.

[124] Camma C, Petta S, Enea M, Bruno R, Bronte F, Capursi V, et al. Cost-effectiveness of boceprevir or telaprevir for untreated patients with genotype 1 chronic hepatitis C. Hepatology 2012;56(3):850–60.

[125] Liu S, Cipriano LE, Holodniy M, Owens DK, Goldhaber-Fiebert JD. New protease inhibitors for the treatment of chronic hepatitis C: a cost-effectiveness analysis. Ann Intern Med 2012;156:279–90.

[126] Poordad F, McCone J, Bacon BR, Bruno S, Manns M, Sulkowski M, et al. Boceprevir combined with peginterferon alfa-2b/ribavirin for treatment-naive patients with HCV genotype 1: SPRINT-2 final results. Hepatology 2010;52 [abstract LB4].

[127] Suzuki F, Suzuki Y, Akuta N, Sezaki H, Hirakawa M, Kawamura Y, et al. Influence of ITPA polymorphisms on decreases of hemoglobin during treatment with pegylated interferon, ribavirin, and telaprevir. Hepatology 2011;53:415–21.

[128] Thompson AJ, Muir AJ, Sulkowski MS, Patel K, Tillmann HL, Clark PJ, et al. Hepatitis C trials that combine investigational agents with pegylated interferon should be stratified by interleukin-28B genotype. Hepatology 2010;52:2243–4.

[129] Afdhal NH, McHutchison JG, Zeuzem S, Mangia A, Pawlotsky JM, Murray JS, et al. Hepatitis C pharmacogenetics: state of the art in 2010. Hepatology 2011;53:336–45.

[130] Maxmen A. Pharmacogenomics: playing the odds. Nature 2011;474:sS9–10.

[131] Pol S, Aerssens J, Zeuzem S, Andreone P, Lawitz EJ, Roberts S, et al. Similar SVR rates in IL28B CC, CT or TT prior relapser, partial- or null-responder patients treated with telaprevir/peginterferon/ribavirin: retrospective analysis of the realize study. J Hepatol 2011;54:S6 [A13].

This page is intentionally left blank

Chapter 38

Pharmacogenomics of Antifungal Agents

H.R. Ashbee[a] and **M.H. Gilleece**[b]

[a]*Mycology Reference Centre, Department of Microbiology, Leeds Teaching Hospitals NHS Trust, UK,*

[b]*Department of Haematology, St James' Institute of Oncology, Leeds Teaching Hospitals NHS Trust, UK*

Chapter Outline

38.1 Overview of Antifungal Therapy 879
  38.1.1 Polyenes 879
  38.1.2 Triazoles 881
  38.1.3 Echinocandins 884
  38.1.4 Flucytosine 885
  38.1.5 Terbinafine 885
38.2 Public Health Impact of Fungal Diseases 886
38.3 The Need for Pharmacogenomics and Stratified Medicine 886
38.4 Genetic Variability Associated with Antifungal Drugs 887
  38.4.1 Absorption 887
  38.4.2 Distribution 890
  38.4.3 Metabolism 892
  38.4.4 Elimination 895
38.5 Current Successes in Stratified Medicine 895
38.6 Future Perspectives 896
References 896

38.1 OVERVIEW OF ANTIFUNGAL THERAPY

In 1939 the war against fungal diseases began in earnest with the discovery of griseofulvin, a metabolite of *Penicillium griseofulvum* [1]. It was the first truly effective antifungal agent and remains the sole licensed orally active drug for childhood dermatophyte infections. Griseofulvin inhibits fungal mitosis and, while it has good activity against dermatophyte fungi, it has no other clinically useful antifungal activity. The shared eukaryotic origin of fungal and human cells had presented a problem for rational drug development since the common metabolic pathways and cellular organelles increase the risk of dose-limiting toxicity (see Figure 38.1). Previous attempts at therapy, using potassium iodide and the sulphonamides, therefore had limited efficacy.

A decade later, the first of the polyene drugs, nystatin, was discovered, followed in 1955 by amphotericin. Amphotericin B is synthesized by an aerobic actinomycete, *Streptomyces nodosus*. The drug was discovered during a survey of soil samples carried out by a pharmaceutical company looking for antibiotic-producing *Streptomyces* strains. The strain producing amphotericin B was first recovered from the Orinoco River valley in Venezuela. These drugs established the modern era in antifungal therapy and remain widely used today. In addition to the polyenes, two other drug classes dominate systemic antifungal therapy: the triazoles and the echinocandins. Other drugs, such as flucytosine and terbinafine have a role in specific diseases or superficial infections, respectively. A general overview of the spectrum of activity of the currently used antifungal drugs is shown in Table 38.1.

38.1.1 Polyenes

The polyene antifungals that are in current clinical use are nystatin, natamycin, and amphotericin B. They are fungicidal, and their mode of action is disruption of the fungal cell membrane by binding to ergosterol, the main sterol in the membrane. This results in pore formation and leakage of cellular cations and anions, leading to fungal cell death [2].

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00038-4
© 2014 Elsevier Inc. All rights reserved.

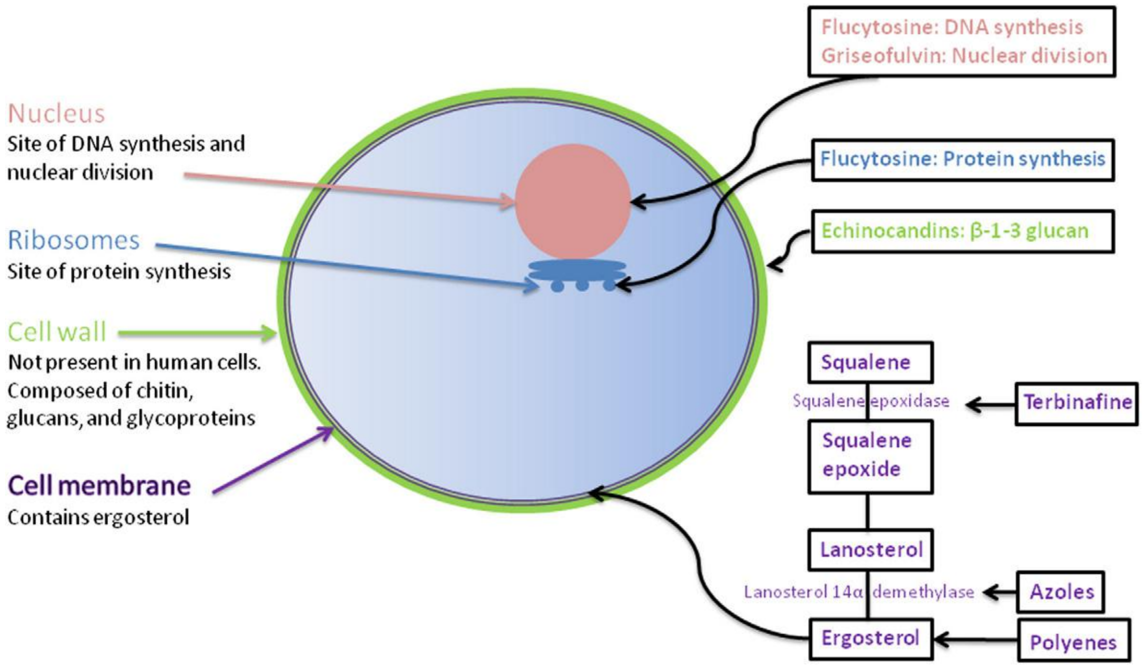

Nucleus
Site of DNA synthesis and nuclear division
Ribosomes
Site of protein synthesis
Cell wall
Not present in human cells. Composed of chitin, glucans, and glycoproteins
Cell membrane
Contains ergosterol
Flucytosine: DNA synthesis
Griseofulvin: Nuclear division
Flucytosine: Protein synthesis
Echinocandins: β-1-3 glucan
Squalene
Squalene epoxidase
Terbinafine
Squalene epoxide
Lanosterol
Lanosterol 14α demethylase
Azoles
Ergosterol
Polyenes

FIGURE 38.1 Cellular targets of antifungal drugs.

TABLE 38.1 Overview of *In Vitro* Spectrum of Antifungal Agents

	Amphotericin B	Flucytosine	Fluconazole	Itraconazole	Voriconazole	Posaconazole	Echinocandins
Aspergillus	✓ (except *A. terreus*)	✗	✗	✓	✓	✓	✓
Candida	✓	✓	✓ (except *C. krusei*; may be reduced for *C. glabrata*)	✓ (may be limited for *C. krusei* and *C. glabrata*)	✓	✓	✓
Cryptococcus	✓	✓	✓	✓	✓	✓	✗
Endemic/dimorphic fungi	✓	✗	✓	✓	✓	✓	Limited activity
Fusarium	Some activity	✗	✗	✗	✓	✓	✗
Scedosporium	Not *S. prolificans*	✗	✗	Limited activity	Limited for *S. prolificans*	Limited for *S. prolificans*	✗
Mucoraceous molds	Some activity	✗	✗	✗	✗	✓	✗

Source: Data Compiled from Sabetelli et al. [16], Wagner et al. [34], and Drew [38].

Although the presence of ergosterol in the fungal cell membrane confers some selectivity to the action of polyenes, there are notable toxicities associated with their systemic use. Nystatin and natamycin are currently only used topically and, although initial studies with a liposomal formulation of nystatin (Nyotran) demonstrated good efficacy, they were halted due to infusion-related toxicity. Amphotericin B is used systemically, but as it is

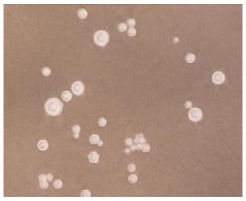
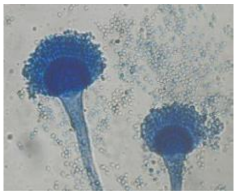
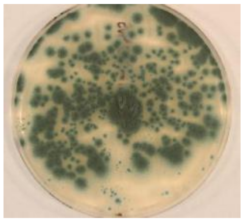
Yeasts Unicellular organisms, which usually reproduce by budding	**Dimorphic fungi** Exist as either yeast or mold forms and the change in form is usually related to temperature	**Molds** Multicellular organisms, with hyphae, and they often produce specific spore-producing structures
Microscopic appearance (x 400 magnification) 	Microscopic and cultural appearance depends on temperature: usually mold at 27ºC and yeast or yeast-like at 37ºC	Microscopic appearance (x 400 magnification) 
Culture appearance 		Culture appearance 
Examples *Candida* *Cryptococcus* *Malassezia* *Trichosporon*	**Examples** *Histoplasma* *Coccidioides* *Blastomyces* *Paracoccidioides*	**Examples** *Aspergillus* *Fusarium* *Scedosporium* Mucoraceous molds Dermatophytes

FIGURE 38.2 Different fungal morphologies with culture and microscopic appearance.

not absorbed orally, it must be administered intravenously (IV). It has a broad spectrum of activity against yeasts and molds (see Figure 38.2) and has been widely used over several decades to treat a range of diseases, including invasive aspergillosis, cryptococcosis, candidosis, the endemic mycoses, and leishmaniasis [2].

In vivo amphotericin B is largely protein bound, with high concentrations in the liver and spleen, and lower concentrations in the lung and kidney. It is eliminated via the urine (21%) and bile (43%), but there is little metabolic breakdown and most of the drug is excreted unchanged over a period of several weeks. The half-life is initially 10–24 hours but after prolonged use may be as long as 15 days [3].

Toxicity associated with amphotericin B includes infusion-related nausea, vomiting, and rigors; chronic toxicity is primarily renal [4]. This toxicity may lead to discontinuation of therapy; in response, lipid formulations were developed and are characterized by reduced toxicity [5]. There are two such formulations currently available in the United Kingdom: Ambisome, which uses liposomes to package the amphotericin, and Abelcet, which uses amphotericin in lipid complexes. These lipid preparations of amphotericin B are also only very slowly eliminated, with only 9% excreted at the end of 7 days [6]. Despite the significantly increased cost, the improved tolerability associated with lipid formulations of amphotericin B has led to a major shift away from the use of amphotericin B deoxycholate except in some resource-poor settings [7].

Drug interactions associated with amphotericin B generally reflect its nephrotoxicity, being most pronounced when concomitant medications undergo renal excretion. Amphotericin B enhances renal toxicity of ciclosporin, aminoglycosides, and some cytotoxic agents such as cisplatin, and it requires vigilant monitoring of renal function if coadministered [2]. Lipid formulations of amphotericin B require no dose adjustment for renal insufficiency and generally none for hepatic insufficiency, although specific recommendations vary with different lipid forms.

38.1.2 Triazoles

The first description of antifungal properties in azole compounds was in 1944 when researchers studying biotin deficiency in animals noted that benzimidazole, which is structurally similar to biotin, had activity against yeasts [8]. This finding was not followed up at the time, but later groups routinely screened other azoles for antifungal activity. Early agents to be developed in the group, including clotrimazole, ketoconazole, and miconazole, are generally now confined to topical use because of unfavorable safety profiles.

Further refinements and chemical modification of these compounds resulted in the triazole group, which is characterized by a five-member azole ring. Triazoles such as fluconazole, itraconazole, voriconazole, and posaconazole are widely used for prophylaxis and treatment of invasive fungal disease (IFD) [9]. The azoles share a common mode of action: inhibition of the fungal cytochrome P450 enzyme (CYP450) lanosterol 14-α-demethylase, which facilitates conversion of lanosterol to ergosterol, a component of the fungal, but not the human, cell membrane. Disruption of ergosterol biosynthesis causes damage to the fungal cell membrane, accumulation of methylsterols, and either cell lysis leading to fungal cell death or inhibition of fungal cell growth, depending on the azole and fungal species [9].

All classes of azoles are substrates and inhibitors of CYP450 enzymes, but the extent to which they are metabolized by or inhibit the various CYP isoforms varies between drugs [10–12]. Individual triazoles exhibit different properties, spectra of activity, and therapeutic indications, and these are discussed next. Their pharmacokinetic properties are summarized in Table 38.2.

38.1.2.1 Fluconazole

The first triazole antifungal to be marketed appeared in 1990. This was fluconazole, a fluorinated bistriazole (i.e., with two triazole rings). Available in capsule and oral solution, and as an intravenous preparation, it has activity against *Cryptococcus*, most *Candida* species (except *C. krusei* and sometimes reduced activity against *C. glabrata*), and some dimorphic fungi. It is generally considered to be fungistatic. Because of its hydrophilic properties, it has excellent oral bioavailability (85–90%) and its absorption is unaffected by gastric pH [9,13]. As fluconazole is largely excreted via the kidneys, dose adjustment is required in patients with reduced creatinine clearance, and it should be used with caution in patients with hepatic insufficiency, although the data are limited. Fluconazole has a favorable safety profile and minimal toxicity, although hepatotoxicity and prolongation of the QT interval in ECG have been noted [3].

Fluconazole is a substrate of CYP3A4 and inhibits CYP 2C9, 2C19, and, to a lesser extent, CYP3A4, an effect that may only occur at doses >200 mg/d [10]. It is also a substrate of P-glycoprotein (P-gp), but not an inhibitor [14]. Drug interactions mediated via the CYP450 enzymes are numerous and include those with alfentanil, ciclosporin A, and midazolam mediated via interactions with CYP3A4; those with amitriptyline, fluvastatin, and phenytoin are mediated via inhibition of CYP2C9 [15].

38.1.2.2 Itraconazole

Itraconazole is a lipophilic triazole, available as capsules, oral solution, and, in some countries, a concentrate for parenteral administration. It has a broad spectrum of activity,

TABLE 38.2 Pharmacokinetic Properties of Triazole Antifungals

	Fluconazole	Itraconazole (Solution)	Voriconazole	Posaconazole
Oral bioavailability (%)	>90 (no food effect)	~50 (fasting)	96 (Fasting)	8–48 (with food)
Protein binding (%)	12	>99	58	98–99
Tissue distribution (%) (as a percentage of serum concentration if known)	Eye: 28–75; CSF: 50–90; bronchial secretions and saliva: 90–100	High: lung, spleen, kidney, liver; moderate: bronchial secretions; very low: eye, CSF, saliva	High: lung, spleen, kidney, liver; moderate: brain, eye, CSF: 22–100	Moderate: bronchial secretions/sputum; low: CSF
Metabolism	<20% metabolized; minor hepatic (CYP3A4)	Extensive hepatic (CYP3A4); many metabolites (including microbiologically active hydroxy-itraconazole)	Oxidative hepatic (mainly CYP2C19; less CYP3A4 and CYP2C9); inactive metabolites	Minimal metabolism; oxidation, glucuronidation (CYP3A4)
Half-life (hour)	25–30	24	6	25
Adult dosing regimen	Varies depending on indication	Varies depending on indication	Maintenance dose of 200 mg twice daily (for patients > 40 kg)	400 mg twice daily or 200 mg 4 times daily
Route of elimination	80% unchanged in urine	3% as metabolites (3–18% unchanged) in urine; 54% as metabolites in feces	80% as metabolites in urine; 20% as metabolites in feces	14% as metabolites in urine; 66% unchanged in feces

Source: Data from Several Sources. See [3,13,32,94–96,125].

including *Candida, Aspergillus,* dimorphic fungi, and some melanized fungi, but not *Fusarium, Scedosporium*, or the mucoraceous molds [16]. Its activity may be either fungistatic or fungicidal depending on the fungal species [17]. Oral bioavailability of itraconazole varies with formulation. Absorption of the capsules is dependent on gastric pH, with increased absorption at acidic pH, while the solution is best absorbed during fasting [9]. Itraconazole has nonlinear pharmacokinetics and because of variable absorption, drug concentrations show significant variability in individuals [18].

In patients with renal insufficiency, use of the IV formulation is not recommended if creatinine clearance is less than 30 ml/min, as the cyclodextrin vehicle is cleared by glomerular filtration. In hepatic insufficiency, there is an increase in the elimination half-life and dose modification may be required [18].

Adverse events associated with itraconazole use include nausea and vomiting (particularly for the oral solution), hepatotoxicity, rash, and headache [18]. There is some evidence that the occurrence of these toxicities correlates with the itraconazole concentration in serum.

Itraconazole is a substrate and potent inhibitor of CYP 3A4, with many metabolites also showing potent CYP3A4 inhibitory activity [19]. It is also an inhibitor of CYP 2C9 and P-gp [15]. These interactions form the basis for many drug interactions seen with itraconazole. Interactions mediated via CYP3A4 include those with fluticasone, lovastatin, phenytoin, rifampin, terfenadine, triazolam, and vincristine; those mediated via P-gp include vincristine, cimetidine, and digoxin [15].

38.1.2.3 Voriconazole

Voriconazole is a second-generation triazole based chemically on fluconazole but with a much broader spectrum of activity. It is active against *Candida* (including most of the fluconazole-resistant species), *Cryptococcus*, and many species of *Aspergillus*; it has some activity against *Fusarium* [16]. Voriconazole activity may be either fungicidal or fungistatic depending on the fungal species [17]. It is the treatment of choice for invasive aspergillosis and is used to treat infections due to *Scedosporium* or *Fusarium*. Formulations of voriconazole include tablets, an oral suspension and an IV solution.

The suspension has good oral bioavailability, especially during fasting, but this is reduced in the presence of high-fat meals [20]. The pharmacokinetics of voriconazole are nonlinear in adults, meaning that relatively small changes in dosing can result in unpredictable changes in drug exposure—being either much larger or smaller than expected.

Voriconazole is extensively metabolized by the CYP 450 enzymes—mainly CYP2C19 but also CYP2C9 and CYP3A4 [21]. It is eliminated as inactive metabolites predominantly via urine (80%), with about 20% excreted in the feces [22].

The vehicle used for IV voriconazole (sulphobutylether beta cyclodextrin sodium, or SBECD) is cleared by the kidneys and hence can accumulate in mild–moderate renal dysfunction; in these patients, the oral formulation is preferable. In mild–moderate hepatic insufficiency, the standard loading dose should be used and then half the maintenance dose, but there are no data for use in severe hepatic insufficiency.

Toxicity associated with voriconazole includes visual disturbances, rash, and hepatotoxicity. Visual disturbances often decrease with continued use and may manifest as photophobia, color changes, blurred vision, or flashes of light; they seem to correlate with high drug concentrations [23]. Cutaneous adverse events are often associated with sun exposure and usually resolve after discontinuation of the drug, although an increased risk of squamous cell carcinoma has been reported with voriconazole use in solid-organ transplant recipients [24]. Hepatotoxicity includes elevated bilirubin, alkaline phosphatase, and transaminases, and correlates with drug concentrations [25].

Voriconazole is both a substrate and an inhibitor for CYP2C19, 2C9, and 3A4, and is also an inhibitor of 2B6 [26]. These properties underpin its multiple interactions with drugs such as sirolimus, alfentanil, cyclosporine, diclofenac, methadone, efavirenz, and midazolam [15].

38.1.2.4 Posaconazole

Posaconazole is the latest triazole to come to market. It was developed from itraconazole and is available only as an oral suspension, although a new tablet formulation is currently under development [27]. Posaconazole has the broadest spectrum of activity *in vitro* of all the azoles, with good activity against *Candida, Cryptococcus, Aspergillus*, many dimorphic or melanized fungi, and several of the mucoraceous molds [16]. It is fungicidal against *Aspergillus, Cryptococcus*, and *Candida* [28]. Oral bioavailability is very dependent on gastric pH, with absorption maximized in the presence of fatty foods [29]. As with itraconazole, absorption is improved by concomitant administration of acidic drinks, but it is saturated at a dose of 800 mg/d and improved by dividing the dose. Drug concentrations achieved in serum vary between patient groups and are known to be reduced in the presence of mucositis and diarrhea, or with the use of proton pump inhibitors (PPIs) or H_2 antagonists [30].

Most of the metabolism of posaconazole is by glucuronidation, with only a limited amount via oxidative CYP450-mediated transformation [31]. Posaconazole has a long half-life and is mainly eliminated unchanged in the feces (66%), with some excreted via the urine, mainly as metabolites (14%) [32]. It can be used in renal insufficiency with no dose modification, and limited data suggest that it should be used with caution in hepatic insufficiency. Toxicity associated with posaconazole use is generally mild, with headache, gastrointestinal upset and rash reported, although at the moment there are no data to suggest that these are related to the drug's concentration. Although posaconazole is minimally metabolized by the CYP450 system, it is an inhibitor of CYP3A4 and P-gp, resulting in some drug interactions, notably with rifabutin, phenytoin, efavirenz, and fosamprenavir [15].

38.1.3 Echinocandins

The echinocandin antifungals, the first new class of antifungals for 15 years, are synthetically modified lipopeptides with three currently licensed: caspofungin, micafungin, and anidulafungin. The lead compound for anidulafungin was discovered in 1974, and all three drugs are derived from fermentation products of fungi [33]. Echinocandins are noncompetitive inhibitors of β-1-3-glucan synthase, an enzyme critical to the synthesis of β-1-3-glucan, which is a major component of the fungal cell wall. The absence of β-1-3-glucan from the cell wall causes loss of rigidity and integrity with resultant lysis. Echinocandins target the *FKS1* gene, which encodes a component of the β-1-3-glucan synthase. Since mammalian cells do not possess a cell wall or β-1-3-glucan synthase, the action of echinocandins is specific to fungi, probably explaining their excellent safety profile. Parenteral administration is obligatory, as their large lipophilic molecular structures have very low oral bioavailability (<10%) [34]. The three echinocandins all have a similar spectrum of activity, a reflection of their mode of action, and they are active against *Candida* and *Aspergillus,* but not against *Cryptococcus* and mucoraceous molds which have less β-1-3-glucan in their cell walls [33,34]. The echinocandins are fungicidal for *Candida* and fungistatic for *Aspergillus* [34]. The pharmacokinetic properties of the echinocandins are summarized in Table 38.3.

Use of the echinocandins is generally well tolerated, although histamine release, some liver toxicity, and headache have all been reported. Caspofungin use is relatively commonly associated with fever and local irritation at the site of administration, while rashes can occur infrequently with all three drugs [35].

The echinocandins are poor substrates for the CYP450 enzymes, are not substrates for P-glycoprotein, and have few drug interactions. Caspofungin is an inhibitor of CYP3A4

TABLE 38.3 Pharmacokinetic Properties of the Echinocandins

	Caspofungin	Anidulafungin	Micafungin
Route of administration	Intravenous only	Intravenous only	Intravenous only
Protein binding (%)	96	>99	99.8
Metabolism	Hydrolysis and N-acetylation	>90% degradation in blood; no hepatic metabolism	>90% degradation; hepatic metabolism; limited role for CYP450
Tissue distribution	High: kidney, liver; moderate: lung, spleen; low: brain, eye	High: lung, liver Moderate: spleen, kidney Low: brain, eye, CSF	High: lung, kidney, liver; low: eye
Half-life for β phase (hour)	11	36–52	11–17
Adult dosing regimen	70 mg loading, then 50 mg if <80 kg or 70 mg if >80 kg	200 mg loading dose, then 100 mg/day	100 mg/day (if >40 kg) or 2 mg/kg/day if <40 kg for invasive candidosis
Route of elimination	41% urine, 35% feces—mostly as metabolites in both	Urine <10%, >90% feces—mainly as metabolites in both	<1% urine; ~70% feces
Dose adjustment in renal insufficiency	None	None	None
Dose adjustment in hepatic insufficiency	Mild: none Moderate: 70 mg loading dose then 35 mg daily Severe: use with caution	Mild–severe: none	Mild–moderate: none Severe: not recommended

Source: Data from Bellman [3], Johnson and Mohr [33], and Wagner et al. [34].

[36]. While ciclosporin A and tacrolimus do interact with it, there is no requirement for dose adjustment although liver function and tacrolimus levels respectively should be closely monitored.

Coadministration of powerful hepatic enzyme inducers (e.g., rifampicin, efavirenz, phenytoin, carbamazepine) require an increase in the daily dose of caspofungin to 70 mg/d [35]. The mechanism of this is unknown, but it is perhaps surprising given that caspofungin is not metabolized by the CYP enzymes. It may be a nonspecific effect or could perhaps be due to induction of the N-acetyltransferases that are involved to some extent in caspofungin metabolism.

Micafungin administration with sirolimus or nifedipine may increase exposure to those drugs and so dosage adjustments may be required if toxicity occurs. Anidulafungin has as yet no known clinically relevant drug interactions [33].

38.1.4 Flucytosine

Flucytosine is a fluorinated pyrimidine that was originally developed as an agent to treat leukemia. It has two modes of action: inhibition of protein synthesis and inhibition of DNA synthesis. The drug penetrates the fungal cell wall via cytosine permease and, once in the cell, is deaminated to form 5-fluorouracil. This is incorporated into fungal RNA instead of uracil and hence disrupts protein synthesis. DNA synthesis inhibition occurs via conversion of 5-fluorouracil to 5-fluorodeoxyuridylic acid, a noncompetitive inhibitor of thymidylate synthase [37]. Flucytosine has fungistatic activity against *Candida* and *Cryptococcus*, and its main current indication is for the treatment of cryptococcal meningitis, in combination with amphotericin B, with which it shows synergy, [38] or the azoles [37]. Flucytosine is available in oral and IV preparations, although the oral preparation is not available in all countries.

Flucytosine has good oral bioavailability, penetrates well into body fluids because of its low protein binding, and achieves good concentrations in cerebrospinal fluid, bone, spleen, liver, heart, and lung. It is mainly eliminated via the kidneys, with almost all of the dose eliminated by glomerular filtration, mainly as unchanged drug. The half-life in patients with normal renal function is 3–4 hours, but can be up to 85 hours in patients with severe renal failure. Renal insufficiency necessitates dose alterations, as does renal replacement therapy, which removes a significant proportion of the drug [37,38]; however, there are no recommendations for use in hepatic insufficiency.

Use of flucytosine is associated with gastrointestinal, hepatic, and hematological toxicities. These side effects may reflect the conversion of flucytosine to 5-fluorouracil [39] and correlate with the levels of flucytosine in serum, with levels above 100 mg/L associated with leukopenia and thrombocytopenia. Hematological toxicity is more common in patients with advanced HIV infection [38]. Therapeutic drug monitoring is required to ensure appropriate dosing and to prevent toxicity.

A particular problem when flucytosine is used as a single agent is the rapid development of resistance [40], so combined use with amphotericin B (or its lipid formulations) is now recommended. Flucytosine has limited drug interactions, but is contraindicated with antiviral nucleoside drugs such as ganciclovir. It may interact with phenytoin and cytarabine, although the evidence for the latter is sparse [38].

38.1.5 Terbinafine

Terbinafine is a member of the allylamine group of antifungals. It inhibits ergosterol biosynthesis by inhibiting squalene epoxidase, a catalyst of the conversion of squalene to squalene epoxide. This resulting accumulation of squalene within the fungal cell and depletion of ergosterol in the fungal cell membrane causes cell death [41]. Terbinafine is highly selective for fungal squalene epoxidase and has no effect on human sterol synthesis [42]. It has *in vitro* activity against a wide range of organisms, including dermatophytes, several melanized molds, *Aspergillus*, *Cryptococcus*, and several dimorphic fungi, although it is mainly used to treat dermatophytosis [43]. It is available as oral and topical preparations.

Terbinafine has good oral bioavailability (70–80%), and absorption is not affected by food. All allylamines are highly lipophilic and hence terbinafine achieves high concentrations in the stratum corneum, sebum, and nail. It is extensively metabolized in the liver, and 15 metabolites have been identified which are subsequently excreted in urine (80%) and feces (20%) [44]. At least 7 CYP enzymes are involved in terbinafine metabolism, of which 1A2, 2C9, and 3A4 are the most important [45]. Elimination takes place slowly, with an initial half-life of 22 hours and another at 90 hours, probably reflecting slow release from adipose tissue [44]. Use of terbinafine is not recommended in patients with active or chronic liver disease or with creatinine clearance <50 ml/min. It has been associated with gastrointestinal effects, skin eruptions, alteration or loss of taste, and hepatobiliary disturbances [46].

Initially, it was thought that terbinafine did not have any inhibitory effect on the CYP enzymes [47], but it is now known to be a potent competitive inhibitor of CYP2D6, although it has been suggested that one of its metabolites, rather than the parent drug, causes this [48]. Because CYP2D6 constitutes only 1.5% of the total liver CYP, there are minimal interactions via this route, except with venlafaxine [49] and amitriptyline [50]. Cimetidine decreases clearance of terbinafine, while fluconazole coadministration causes increased exposure to terbinafine via inhibition of CYP2C9 and 3A4.

38.2 PUBLIC HEALTH IMPACT OF FUNGAL DISEASES

The public health impact of fungal diseases is significant. The impact of superficial versus systemic disease is important for different reasons. Superficial fungal infections are extremely common: various studies estimated that 20–40% of adults in developed countries have some form of fungal foot disease [51], there is an increasing incidence of tinea capitis in children [52], and superficial candidosis is the most commonly seen manifestation of candidiosis [53]. From a public health viewpoint, superficial infections are important because those caused by dermatophytes are infectious and are readily transmitted via communal facilities, such as sports centers or within the family home. While skin infections are not associated with any significant morbidity, the breaches in skin integrity they cause may become a portal of entry for other organisms that may cause more serious infections such as cellulitis. Dermatophytosis may result in alopecia or permanent scarring from scalp infections and difficulty walking in elderly patients with grossly thickened or clawed toenails due to fungal nail infection.

Additionally, because of the sheer numbers of patients affected, the potential cost of medical consultations, diagnosis, and treatment for fungal foot disease alone is huge. Figures from the *Prescription Cost Analysis for England* for 2012 show that prescription costs for terbinafine and griseofulvin, the main drugs used to treat dermatophytosis, totaled £3.8 and £1.2 million, respectively, highlighting the significant cost of treatment alone (www.hscic.gov.uk/catalogue/PUB10610).

For IFD, the public health impact is also significant, despite the small numbers of people affected. Patients in whom systemic diseases are likely to develop are those who are immunocompromised, either because of underlying disease (e.g., leukemia, AIDS) or because of treatments for their disease (e.g., cytotoxic agents, immunosuppressives, organ or haematopoietic stem cell transplantation (HSCT)). Multiple studies in many countries have shown that the incidence of IFD is increasing, partly because more patients are surviving their underlying diseases and partly because more aggressive treatments are being used, resulting in many profoundly immunosuppressed, at-risk individuals [54].

Systemic diseases are relatively infrequent when compared to the prevalence in the community of superficial fungal infections. However, the cost associated with each episode of IFD is disproportionately high. This is because IFDs are associated with increased length of hospital stay and prolonged courses of antifungal drugs. A recent American study estimated that a patient with a systemic fungal infection stays in the hospital an extra 7 days and costs an extra $29,000 compared to a patient with the same underlying disease but no fungal infection [55]. Overall, this represents an additional cost of $1.89 billion, demonstrating that the cost associated with IFD has a significant public health impact.

The recent emergence of *Aspergillus fumigatus* with resistance to the azoles has been documented. Although there is a low level of resistance in many countries throughout Europe, the levels in Holland and the United Kingdom are higher. In Holland, the widespread use of azole-based compounds in agriculture has been implicated as the cause [56], but in the United Kingdom most of the resistance has occurred in patients treated with azoles, usually for months or years, for chronic aspergillosis [57]. The development of resistance to an important class of drugs, which often confers cross-resistance between members of the class, is of concern and may impact our ability to treat these infections in the future.

38.3 THE NEED FOR PHARMACOGENOMICS AND STRATIFIED MEDICINE

Treatment of fungal disease has been increasingly successful over the last 10–20 years, partly due to the advent of better antifungal drugs with broader spectra and less toxicity; and partly due to improvements in diagnosis of IFD, enabling earlier interventions. However, the morbidity and mortality associated with some fungal diseases is still very high.

Various risk factors predispose to IFD, including malignancy (particularly hematological malignancy), the use of steroids or antimicrobial agents, surgery, burns, uncontrolled diabetes, AIDS, chemotherapy, immunosuppressive agents, and neutropenia. The risk of IFD can be crudely predicted based on the level of patient immunocompromise. Therefore, patients who have undergone allogeneic HSCT have a greater risk of fungal disease than do patients receiving autologous transplants, which partly relates to their degree of immune suppression [58]. However, even within groups known to be particularly "at risk," susceptibility to fungal disease varies.

Approaches to prevention and diagnosis of fungal disease vary among institutions. Some use antifungal prophylaxis in high-risk patients in an attempt to avoid the disease. The triazole antifungals have been widely used in this setting, with fluconazole, itraconazole, and posaconazole all having a licence for this indication in the United Kingdom. Despite not having a licence for prophylaxis, voriconazole has also been used. The disadvantages associated with this approach include the cost of the drugs, the potential for drug interactions and toxicity, and the inability to achieve appropriate therapeutic levels in all patients. This can then require the use of therapeutic drug monitoring to ensure drug concentrations high enough but not toxic, which incurs a further cost [59]. Another disadvantage is that evidence suggests that the use of mold-active prophylaxis reduces the sensitivity of the *Aspergillus* galactomannan (GM) ELISA, which is widely used in the diagnosis of invasive aspergillosis [60]. This may then lead to the use of empiric antifungal therapy because of a lack of confidence in the results of the diagnostic test.

Another approach is the routine screening of patients with PCR and GM ELISA—a so called "enhanced diagnostic" approach. This relies on the high negative predictive value of the tests and, unless other features of fungal disease occur, it enables the patient to remain on prophylaxis rather than receive empiric antifungal therapy. One study showed no difference in morbidity or mortality when empiric therapy was withheld and a reduction in antifungal expenditure of £124,000 using the enhanced diagnostic approach [61].

Diagnosis of superficial fungal disease is relatively simple, with direct microscopy and culture of skin, hair, or nail samples allowing the recovery of the causative agent in most patients. The situation is more challenging in patients with suspected IFDs. Patients are likely to have serious underlying disease, and the development of new signs and/or symptoms, which are often nonspecific, may be difficult to pinpoint. Direct microscopy and culture may be useful if tissue is available, but is often slow and insensitive; the inability to take biopsies and deep tissue samples in patients with coagulopathies further limits the usefulness of this modality. Detection of antibodies to fungi is rarely helpful in patients who are immunocompromised.

The use of fungal antigen detection kits has become increasingly common, and several are available, including cryptococcal antigen, *Aspergillus* GM, *Candida* mannan, and the more general β-D-glucan detection kits. While the sensitivity and specificity of the cryptococcal antigen kit is very high, the corresponding figures for the GM kit vary with patient group. It is highest in the HSCT recipient and lower in solid organ transplant recipients. Other tests such as PCR have been useful for some diseases and in certain laboratories, but the lack of standardization and variable performance means that they are still not used routinely in many centers [62].

Whatever approach is used, it is recognized that the successful outcome of an IFD relies on the early institution of appropriate antifungal therapy and that successful outcomes correlate with early initiation of therapy [63].

Although there are now many efficacious antifungal drugs available, the prevention, diagnosis, and treatment of IFD still presents a major challenge. To date we still do not understand the subtle differences between patients that impact their risk of infection or their response to antifungal drugs. Developments in pharmacogenomics that better enable us to predict how antifungal drugs will behave in a given patient will help us move toward more tailored or personalized management in this area.

38.4 GENETIC VARIABILITY ASSOCIATED WITH ANTIFUNGAL DRUGS

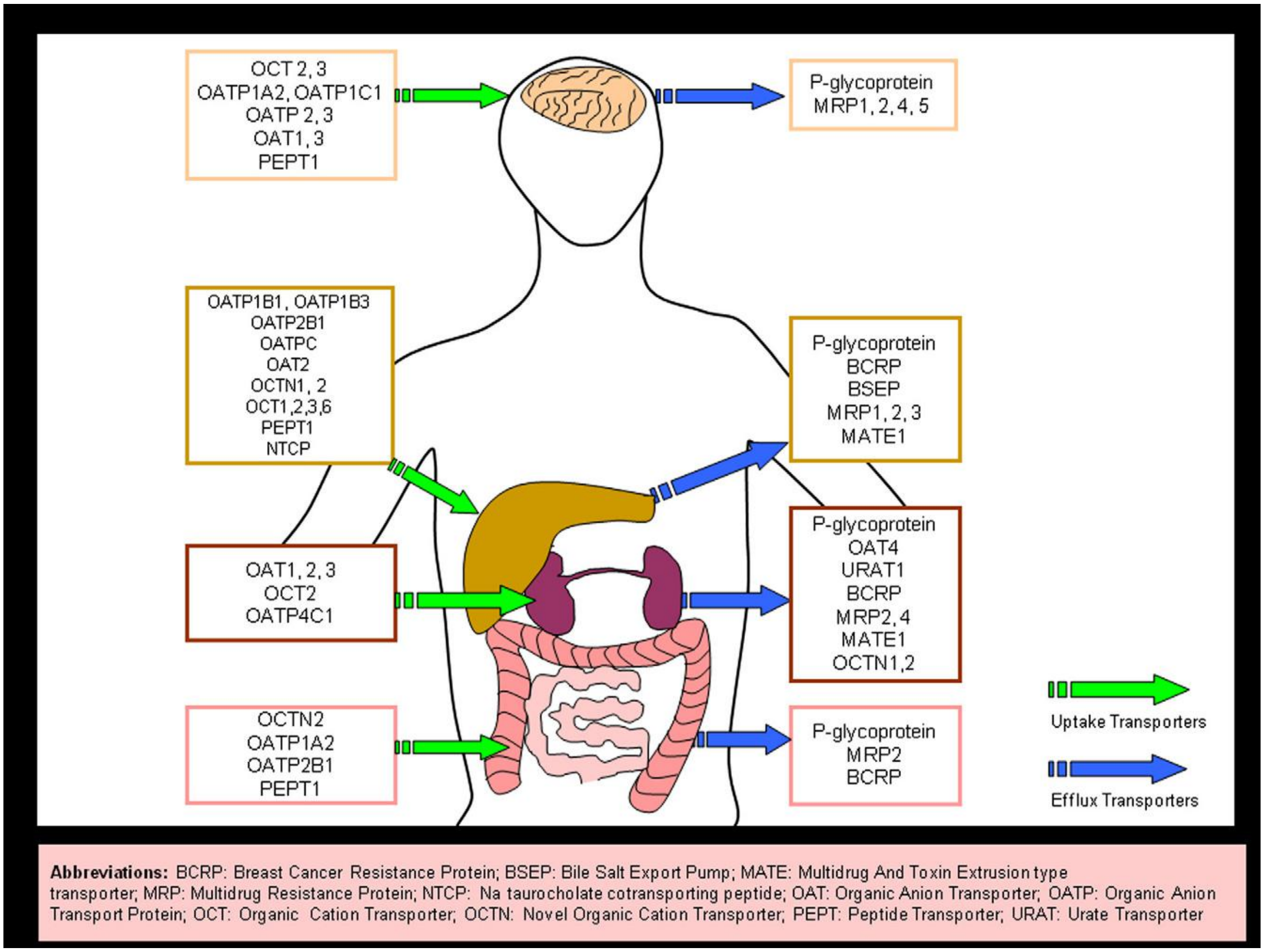
The effect of genetic variability on antifungal drugs may occur at any point during their absorption, distribution, metabolism, or elimination from the body. In addition to genetic variability, there are many other factors that may impact the efficacy of antifungal drugs, including drug interactions and various physiological factors [59,64,65].

38.4.1 Absorption

Antifungal drugs are administered either orally or intravenously, and absorption is affected by many factors but is most variable for drugs that are taken orally. With IV dosing, absorption is assumed to be complete and the total dose reaches the systemic circulation. When the drug is given orally, it must dissolve in the stomach (the extent of which may be affected by gastric pH and the presence and type of food) and may then be subjected to first-pass metabolism in the intestines and liver or transporter-mediated efflux back into the intestinal lumen. These processes are affected by the amount and activity of the drug transporters and metabolizing enzymes at those sites (see Figure 38.3). Biliary excretion at this initial stage may also reduce the net amount of absorption. Because most of the drug is absorbed from the intestines rather than the stomach, any factor that speeds gastric emptying is likely to enhance drug absorption.

Gastric pH is particularly important for the absorption of itraconazole and posaconazole, while it has little effect on the absorption of flucytosine, fluconazole, voriconazole, and terbinafine. Itraconazole and posaconazole are weak bases and hence do not dissolve well in an alkaline environment [66,67]. Gastric pH can vary by up to 10-fold between individuals [68], and men tend to have lower gastric pH than women [69]. Age is now no longer thought to affect pH. Specific patient groups also have alterations in gastric pH due to their disease: those with AIDS have increased gastric pH [70], while critically ill patients show reduced gastric pH [71].

The widespread practice of using PPIs or H_2 antagonists to reduce the risk of gastric ulcers in hospitalized patients [72] will also have a significant impact on gastric pH. Indeed, it has been shown for posaconazole that concomitant PPIs or H_2 antagonists cause a significant reduction in drug exposure in patients [30]; although this is not due to genetic variability, it is an important determinant of drug absorption. For itraconazole capsules (but not oral solution), factors that enhance gastric acidity improve absorption, such as administration with food; whereas factors that decrease gastric acidity (use of PPIs or fasting) reduce absorption [9]. Therefore, patients receiving itraconazole capsules may be affected by variations in gastric pH that


OCT 2, 3
OATP1A2, OATP1C1
OATP 2, 3
OAT1, 3
PEPT1
P-glycoprotein
MRP1, 2, 4, 5
OATP1B1, OATP1B3
OATP2B1
OATPC
OAT2
OCTN1, 2
OCT1,2,3,6
PEPT1
NTCP
P-glycoprotein
BCRP
BSEP
MRP1, 2, 3
MATE1
OAT1, 2, 3
OCT2
OATP4C1
P-glycoprotein
OAT4
URAT1
BCRP
MRP2, 4
MATE1
OCTN1,2
OCTN2
OATP1A2
OATP2B1
PEPT1
P-glycoprotein
MRP2
BCRP
Uptake Transporters
Efflux Transporters

Abbreviations: BCRP: Breast Cancer Resistance Protein; BSEP: Bile Salt Export Pump; MATE: Multidrug And Toxin Extrusion type transporter; MRP: Multidrug Resistance Protein; NTCP: Na taurocholate cotransporting peptide; OAT: Organic Anion Transporter; OATP: Organic Anion Transport Protein; OCT: Organic Cation Transporter; OCTN: Novel Organic Cation Transporter; PEPT: Peptide Transporter; URAT: Urate Transporter

FIGURE 38.3 Major uptake and efflux transporters at different body sites in humans.

result in serum concentrations of itraconazole that are insufficient to prevent or treat their disease.

The main efflux drug transporters in the intestines are P-gp, breast cancer resistance protein (BCRP), and multidrug resistance protein 2 (MRP2) [73]. These transporters are found at the luminal side of enterocytes and are important in limiting the uptake of drugs that are substrates for them. They are all members of the ATP-binding cassette transporters and use ATP to allow them to pump substrates against the concentration gradient. It was recently shown that there is a large amount of inter-individual variation in the quantity of each transporter in the intestines and liver [74] and this might contribute to the variability in drug absorption between different individuals. BCRP and P-gp are present in smaller amounts in the liver than in the gut, while MRP2 is present in smaller amounts in the gut.

All of these efflux transporters have genetic polymorphisms which result in clinically significant changes in substrate binding. P-gp has multiple genetic variants, caused by single-nucleotide polymorphisms (SNPs) that vary with ethnicity and cause clinically relevant differences in, for example, clearance of ciclosporin [75]. SNPs in BCRP and MRP2 also occur and can affect the extent of substrate binding [73].

Several of the orally administered antifungals are either substrates or inhibitors of these transporters (see Table 38.4). Itraconazole, fluconazole, and posaconazole are all P-gp substrates, with itraconazole having the strongest interaction with it; itraconazole and posaconazole are also P-gp inhibitors. Voriconazole binds to CaMdr1P, the yeast homolog of human P-gp, suggesting theoretically that it may be a substrate in humans as well, but as yet there are no data to support this contention [76]. Itraconazole is an inhibitor of BCRP. Based on the chemical similarity between itraconazole and posaconazole and the presence of a long side chain in posaconazole (which enhances inhibitory capacity), it is likely that it too is an inhibitor of BCRP [77]. There are little data on the interaction of terbinafine with these transporters, but it is unlikely to be an inhibitor of

TABLE 38.4 Interactions between Antifungals and Phase I and II Enzymes and Drug Transporters

	CYP450 enzymes		P-glycoprotein		BCRP		
Antifungal	*Substrate*	*Inhibitor*[a]	*Substrate*	*Inhibitor*	*Substrate*	*Inhibitor*	*Other transporters*
Amphotericin B	– [11][b]	– [11]	Possibly if given orally [139]	– [11]	ND	ND	ND
Fluconazole	3A4 [10]	2C9 [30.3 μM] [12]; 2C19 [12.3 μM] [12]; 3A4 [13.1 μM] [11]; 1A2 [31]	++ [14]	– [11,14]	ND	– [77]	UGT2B7/UGT1A4 substrate [118]; UGT2B7 inhibitor [119]; not OAT-P1B1/OATP1B3 inhibitor [93]
Itraconazole	3A4 +++ [19]	3A4 [0.033 μM] [11]; 2C9 [>10 μM] [12]; 2C19 [>10 μM] [12]; 1A2 [31]	+++ [15]	[2 uM] [14] [5–50 uM] [11]	ND	[<1 uM] [77]	Suggested OATP inhibitor via interaction with fexofenadine [91] but *in vitro* study found no inhibition [92]; not OATP1B1/OATP1B3 inhibitor *in vitro* [93]; multidrug resistance protein 3 inhibitor [89]; not NTCP/BSEP inhibitor; UGT1A4 substrate [118]
Voriconazole	2C19 +++ 3A4 ++ 2C9 ++ [21]	2C9 [8.4 μM] [12], [3.62 μM] [26]; 2C19 [8.7 μM] [12], [5.25 μM] [26]; 3A4 [10.5 μM)] [12] [2.9 μM] [26]; 2B6 [1.7 μM] [26]	Possibly [76]	– [140]	ND	– [77]	UGT1A4 substrate [118]; possible UGT2B inhibitor [130]; FMO enzyme substrate [129]
Posaconazole	– [140]	3A4 [1.3 μM)] [31]; 1A2 –; 2C9 –; 2C19 –; 2A6 –; 2D6 –	++ [78]	++ [78]	ND	Likely to be based on long side chain and structural similarity to itraconazole [77]	UGT1A4 substrate [123]
Caspofungin	– [120]	3A4 + [36]	ND	– [98]	ND	ND	NCTP inhibitor; OAT1B1 substrate/ inhibitor [98].
Micafungin	– [12]	2C9 –; 2C19 –; 3A4 [13.5 μM] [11]	– [11]	– [11]	+ [100]	ND	OATP/NTCP (+)/BSEP (++) substrate [100]
Anidulafungin	– [114]	– [102] Not 1A2, 2B6, 2C9, 2C19, 2D6, or 3A inhibitor [114]	ND	ND	ND	ND	Not OATP inhibitor [102]
Flucytosine	ND	ND	ND	ND	ND	ND	
Terbinafine	1A2 +++ 2C9 +++ 3A4 +++ 2C8 +[45]	2D6 [0.2 μM] [45]; 3A4 – [47]	ND	Unlikely [10]	ND	ND	

[a] *IC_{50} (uM) is shown in brackets where figures are available; these may vary between studies in which different probes have been used in vitro.* [b]*Relevant source in reference list.*

ND—no data available; – no inhibitory effect or not a substrate; + weak inhibitor or weak substrate; ++ moderate inhibitor or moderate substrate; +++ potent inhibitor or major substrate.

P-gp because of the lack of effect when it is coadministered with digoxin, a P-gp-dependent substrate [10].

The presence of SNPs in any of the efflux transporters potentially affects the absorption of orally administered antifungals if they cause altered binding due to those polymorphisms. The only drug for which data are available is posaconazole, where the effect of SNP's in P-gp have been studied [78]. Homozygous expression of the T allele results in significantly less expression of P-gp than does expression of heterozygous or wild-type alleles. Volunteers (28 black and 28 Caucasian) were dosed with 400 mg of posaconazole twice daily after food for 8 days, after which concentrations of posaconazole and expression of SNPs in P-gp were investigated. Overall, 41 SNPs were found in the alleles, with 13 individuals being homozygous for the T allele and 43 being either heterozygous or wild-type. There was no difference in drug exposure for either group, and hence it seems that polymorphisms in P-gp have no effect on posaconazole absorption. However, as this study was performed in healthy volunteers, its applicability to patients is uncertain.

There are currently no other data for the other antifungals, but similar studies are of interest particularly for itraconazole, which is a potent substrate and inhibitor of P-gp.

38.4.2 Distribution

Once the drug has been absorbed after oral administration, it must reach the bloodstream. Drugs that are administered intravenously overcome the issues related to absorption, but they too must move from the bloodstream to the interstitial and intracellular fluid and ultimately to the site or sites of the fungal disease. Blood flow to the various organs is one of the main determinants of tissue delivery, with the well-perfused organs (e.g., liver, kidney, brain) receiving the highest concentrations initially, with the skin, muscles, and fat receiving the drug possibly several hours later. The extent to which the drug partitions between blood and tissue is determined by how much the drug is bound by serum proteins, as only an unbound drug can move across cell membranes. The main proteins that bind drugs in serum are albumin—the major carrier for acidic drugs—and α_1 acid glycoprotein (AAG), which binds basic drugs. Lipoproteins also act as carriers for some drugs.

The concentrations that are achieved for different antifungals at different body sites vary considerably (see Tables 38.2 and 38.3). Some antifungals are highly protein bound (e.g., itraconazole, posaconazole), while others have much lower protein binding (e.g., fluconazole, flucytosine). Itraconazole and posaconazole are mainly bound to albumin, although itraconazole does also bind, to a lesser extent, to AAG. Although much less protein bound, fluconazole preferentially binds to AAG [79]. Amphotericin B is bound by lipoproteins, albumin, and AAG [80,81], while terbinafine is strongly bound to serum proteins and the binding is even across the lipoproteins and albumin [44]. All echinocandins are extensively protein bound, mainly to albumin, but the extent of binding of caspofungin varies between healthy volunteers and patients [34].

Many serum proteins are now known to have genetic variants, with human serum albumin having 67 genetic variants [82] and AAG having 2 [83]. These variants have altered substrate binding and because of the extensive protein binding of many antifungals, this may affect their distribution. If protein binding of the antifungal were reduced, this could allow greater concentrations to move into the cells more quickly; conversely, if the binding were increased, this could reduce the rate at which the drug dissociates and moves into the tissues. At the moment, however, this is entirely speculative and there are no data about how antifungals are affected by serum protein genetic variants.

Drugs are taken into cells by either passive diffusion or active transport. As well as the extent of its protein binding, a drug's ability to cross the cell membrane and hence enter the cell is determined by factors such as its size, degree of ionization, and lipid solubility of its ionized and nonionized forms. Passive diffusion into the cell usually occurs because the drug is soluble in the lipid bilayer and its movement occurs down a concentration gradient. Hydrophilic drugs penetrate cell membranes via the aqueous channels. In contrast, active transport requires the expenditure of energy and usually moves a drug against the concentration gradient, is saturable, and may be subject to competitive inhibition. Uptake transporters include those that transport anionic drugs (organic anionic transporters, or OATs) and organic anionic transporting polypeptides (OATPs), as well as those that transport cationic drugs (organic cationic transporters, or OCTs) and peptide transporters (PEPTs). As with efflux transporters, genetic polymorphism occurs for the uptake transporters resulting in alterations in substrate binding [84].

The distribution of drugs around the body and entry into specific tissues is determined partly by the uptake transporters. Antifungal drugs must be taken up into tissues to reach the site of the fungal disease, and in sufficient concentrations to allow them to either inhibit or kill the fungi. If the concentration achieved locally is not sufficient, then although the fungus may appear to be susceptible to the agent *in vitro*, the drug will not be efficacious *in vivo*, resulting in clinical failure.

Amphotericin B achieves high concentrations in the liver and spleen, and moderate concentrations in the lung and kidney. The data on the involvement of drug transporters in the uptake of this drug into tissues are very limited. Amphotericin B potentiates the effects of various cytotoxic drugs. This may be due either to increased cellular permeability or to P-gp inhibition [85]. P-gp is involved in the efflux of drugs from many types of cancer cells, which causes resistance to cytotoxic agents; however, if amphotericin B inhibits P-gp, this may enhance both the level and the effects of cytotoxic agents in cancer cells. In contrast, it was shown that the coadministration of amphotericin B and ciclosporin in rats caused an induction of multidrug resistance gene 1 (*MDR1*) expression and higher levels of P-gp, suggesting that amphotericin B may induce P-gp rather than inhibit it [86]. Studies *in vitro* showed that amphotericin B has no activity as a P-gp inhibitor. For this reason, the extent to which amphotericin B interacts with P-gp is still not clear, and hence its role in the distribution of this drug is also unknown. There are no other data on which drug transporters amphotericin B interacts with, so few conclusions can be drawn about the likely effect of drug transporter polymorphisms on amphotericin B distribution.

Fluconazole has low protein binding and is hydrophilic; hence, it achieves good tissue penetration with high concentrations in cerebrospinal fluid (CSF), saliva, bronchial secretions, and brain [3]. It is not an inhibitor of OATP1B1 or OATP1B3, which are both involved in uptake into the liver. Fluconazole is probably taken up into cells and tissues by passive diffusion and so is unlikely to be affected by genetic variability in uptake transporters.

Itraconazole is found in the highest concentrations in the liver, spleen, kidney, and lung, and in very low concentrations in the eye and CSF. It is known to be a potent substrate and inhibitor of P-gp, which is likely to affect its tissue distribution. Absorption of itraconazole can be variable, and it may be that as a substrate of P-gp, some of the drug is pumped back into the lumen of the intestine by P-gp and hence reduces absorption. Similarly, P-gp acts within the brain to limit drug uptake into brain tissue. In a mouse model, itraconazole was shown to be effluxed from the brain by P-gp; in knockout mice that lacked P-gp, this efflux mechanism was removed, allowing accumulation of itraconazole within the brain [87]. Similarly, an inhibitor of P-gp allowed improved concentrations of itraconazole in a mouse model by overcoming the efflux due to P-gp [88]. The low concentrations of itraconazole seen in the brain and CSF suggest that this may be a possible mechanism responsible in humans, too. In the presence of genetic variants of P-gp, this has the potential to alter both the uptake of itraconazole from the intestine and its access into the brain. Itraconazole is also a substrate or inhibitor of several other drug transporters, including BCRP and MRP3, but at physiological concentrations it does not inhibit the bile salt extract pump (BSEP) or the hepatic uptake transporter Na+ taurocholate-cotransporting peptide (NTCP) [89]. Polymorphisms in BCRP or MDR3 could therefore also impact itraconazole distribution. Based on its interaction with atorvastatin [90] and fexofenadine [91], it has also been suggested that itraconazole may be an inhibitor of OATP1B1, which is important in uptake into the liver, although *in vitro* data do not support this idea [92,93]. Therefore, because of extensive interactions with drug transporters, there are multiple points at which genetic variability could impact the distribution of itraconazole in the body.

Voriconazole is found at high concentrations in the lung, spleen, liver, and kidney, with moderate concentrations in the brain, eye, and CSF [94,95]. The data on voriconazole's interaction with drug transporters are currently very limited. Voriconazole does not inhibit P-gp or BCRP, and there is only circumstantial evidence that it may be a substrate for P-gp. Its penetration into brain and CSF tends to suggest that it is not a substrate of P-gp or it would be subject to efflux from the brain and CSF, as is seen for itraconazole. As more information emerges about interactions between voriconazole and drug transporters, it may be possible to determine whether there is potential for genetic variants to impact its distribution; at the moment, however, there is no obvious point where this might occur.

Little is known about the tissue distribution of posaconazole except that it does not achieve high concentrations in CSF, although it achieves concentrations above the MIC for most fungi in various compartments of the lung [96]. Posaconazole is a substrate for P-gp and hence P-gp-mediated efflux may explain some of the absorption problems seen with this antifungal and its low CSF concentration. Based on its structural similarity to itraconazole, it is probably an inhibitor of BCRP; if so, its inhibition of BCRP in the intestine might enhance its own uptake. Potentially, polymorphisms in either P-gp or BCRP might affect the distribution of posaconazole; at the moment, however, the evidence for this is limited.

Caspofungin is a large molecule and highly protein bound; nevertheless, it achieves high concentrations in the kidney and liver, moderate concentrations in the lung and spleen, but only low concentrations in the brain and eye. Its profile in plasma is determined by its rate of distribution into tissues rather than by its metabolism or elimination [97]. Uptake into the liver is a two-step process. The first part is a rapid reversible absorption to the cell surface, followed by a second slower phase with transport across the cell membrane [97,98]. OATP1B1 mediates uptake into hepatocytes, with no uptake by OATP1B3. Caspofungin has a relatively low affinity for OATP1B1 and is also a modest inhibitor of this transporter [98]. It also inhibits NTCP, OCT1, and OAT1 at 100 μM, although at physiological concentrations of 1–10 μM, this inhibition may not occur. Caspofungin has no inhibitory activity against P-gp [98].

The importance of OATP1 in the transport of caspofungin into the liver is illustrated by the interaction with rifampin, which inhibits OATP1 and causes a decrease in the trough levels of caspofungin by 14–31% after multiple rifampin doses [99]. The most important drug transporter for caspofungin is OATP1B1, and so it is only OATP1B1's genetic variants that may affect caspofungin distribution. Multiple genetic polymorphisms have been described for OATP1B1, including some that occur with relatively high frequency in several ethnic groups and cause reduced transporter activity. Therefore, if any of these polymorphisms occur in patients who are receiving caspofungin, they may potentially affect the drug's uptake into the liver and hence its distribution.

Micafungin accumulates to high concentrations in the liver, lung, and kidney, and its uptake into the liver is well characterized and involves both passive and active transport. The hepatic uptake transporter NTCP is most important in micafungin uptake, with OATP playing a minimal role. This finding is unusual, as it was previously thought that NTCP was a specific transporter solely for bile salts. P-gp has no involvement in transport of micafungin [100]. NTCP shows genetic variability across different ethnic groups, resulting in reduced activity [101]; thus, individuals expressing variants in this transporter may have altered ability to take micafungin into the liver.

In clinical studies, anidulafungin reached high concentrations in the lung and liver, moderate concentrations in the spleen and kidney, and low concentrations in the brain, CSF, and eye. At concentrations achieved at steady-state (1–3 μM), anidulafungin is not a substrate or inhibitor for OATP-B or OATP-C, although at higher concentrations (10 μM) it causes inhibition of OATP1B3 [102]. The potential for anidulafungin distribution to be affected by genetic polymorphisms therefore appears to be very limited.

Flucytosine has very low protein binding and is very hydrophilic. Because of these two factors, it can achieve good tissue concentrations in CSF, saliva, bronchial secretions, and brain [3]. There are currently no data on the interaction of flucytosine with any of the drug transporters, but because it appears to be taken up by passive diffusion into cells it is unlikely to be affected by genetic uptake transporter variability.

Terbinafine is primarily used for the treatment of superficial fungal infections and so its distribution in other tissues of the body has been little studied. The drug is lipophilic and achieves good concentrations in skin, hair, and nail. It is unlikely to be a P-gp inhibitor because of its lack of interaction with digoxin [10]. However, there are no other data on its interaction with drug transporters. Given this lack of information, it is not currently possible to determine whether polymorphisms in drug transporters may affect terbinafine distribution.

38.4.3 Metabolism

Many drugs are lipophilic, which allows them to transit biological membranes but hinders their removal from the body. These compounds must be metabolized to more hydrophilic molecules for excretion. Drug metabolism reactions are classified either as phase I or phase II. Phase I reactions include oxidation, reduction, and hydrolysis and generally result in inactive drugs. Phase II reactions conjugate the result of phase I reactions with a second molecule and result in metabolites with improved water solubility. Phase I reactions are carried out by superfamilies of related enzymes, including CYP450 enzymes and flavin monooxygenases (FMO).

CYP450 enzymes metabolize a wide range of endogenous and exogenous substrates, including sterols, fatty acids, vitamins, and a multitude of drugs. Substrate specificity for each enzyme is often very broad—for example, CYP3A4 is responsible for metabolizing about 50% of clinically used drugs, including many immunosuppressives, statins, benzodiazepines, and other drugs. This lack of reaction specificity means that if coadministered drugs are metabolized by the same enzyme, there is substantial potential for drug interaction [103].

The expression of different CYP enzymes varies between tissues. The main CYPs expressed within the intestine are CYP3A, accounting for approximately 80% of total protein, and 2C9, which contributes ~14% [104]. In the liver, CYP3A4 contributes 29% of the total protein content; CYP2C (2C8, 2C9, 2C18, and 2C19), 18%; CYP 1A2, 13%; and other CYPs, smaller amounts [105]. CYP3A expression levels vary throughout the length of the intestines, with the highest levels in the duodenum and middle jejunum, declining in the distal jejunum and ileum [106]. As well as this variability in CYP enzyme distribution, enzyme activity can vary significantly between individuals, with up to 20-fold variability in the activity of CYP3A4 [107]. Because of this, the extent of metabolic activity on drugs can be markedly different between individuals and hence the concentrations achieved in the tissues, even though the individuals received the same initial dose.

Compounding this variability in expression and lack of enzyme specificity is the widespread occurrence of genetic variants within the different CYP alleles. Many of the alleles have SNPs that result in reduced, lack of, or occasionally increased metabolic activity. CYP3A4/5 currently has over 40 described allelic variants, 2C9 has more than 60, and 2C19 has more than 30 [108]. The prevalence of these alleles varies with ethnicity, as discussed in Chapter 16. For 2C9, about 80% of the Caucasian population express the wild-type allele (2C9*1A) and about 20% express the allele's genetic variants [109]. CYP2C19 has several variants that result in a "poor metabolizer" phenotype, where individuals are not able to clear CYP2C19 substrates as well as those who have wild-type alleles and are either heterozygous or homozygous for the wild-type 2C19*1A allele. The incidence of 2C19 poor metabolizers is about 2–5% in Caucasians; in Asian populations this may be 15–20% [110]. The prevalence of polymorphic variants of 3A4 alleles is very low (about 1% in Caucasians), but it has been suggested that some of the activity ascribed to 3A4 is actually due to 3A5, which has a polymorphic allele frequency of 90% in Caucasians [111].

The other main group of enzymes involved in phase I metabolism are the FMOs. These are involved in the oxygenation of compounds to make them more polar and hence easier to excrete. There are five isoforms of the enzyme, which vary in tissue distribution and activity, and have genetic variants with altered substrate activity [112].

Phase II reactions are carried out by various enzymes, including the glutathione-S-transferases (GSTs), UDP-glucuronosyltransferases (UGTs), and N-acetyltransferases (NAT). These reactions result in the conjugation of a substrate and usually ablate its activity. As with the phase I enzymes, genetic variants have been described that result in changes in the enzymes' metabolic activity [113].

Some data exist on the effect of genetic polymorphisms on the metabolism of various antifungals. The magnitude of the effect mainly depends on the way in which the drug is metabolized. For drugs that either are not metabolized at all or undergo metabolism independent of the phase I and II enzymes, the potential for genetic variants affecting them is minimal. This includes anidulafungin, amphotericin B, flucytosine, and fluconazole.

Anidulafungin is unusual among antifungal drugs in that it undergoes a slow nonenzymatic degradation mainly in the tissues, with no involvement of either the phase I or phase II enzymes. Initially, an inactive metabolite is formed that is thought to undergo further degradation by plasma peptidases [114]. Therefore, there are no points at which genetic variants in phase I or II enzymes can impact this process. Unlike anidulafungin, flucytosine and amphotericin undergo virtually no metabolism and are almost entirely excreted unchanged [6], although very small quantities of metabolites have been described for flucytosine [115]. Because of this lack of metabolism, there is no way in which genetic polymorphisms can affect this process.

However, evidence from drug interactions suggests that amphotericin B may have various effects on CYP450 enzymes. In animals, amphotericin B caused reduced concentrations of hepatic CYP enzymes, and in patients with HIV a reduction in CYP enzymes was also reported [116]. The mechanism is unclear and may have been hepatocyte damage rather than a specific effect on CYP. In contrast, amphotericin B was found to induce CYP3A2 and consequently first-pass metabolism of ciclosporin, which may contribute to decreased levels when amphotericin B and ciclosporin are coadministered [86]. Nothing is known about its ability to interact with the genetic variants of these CYP enzymes.

Fluconazole is highly water soluble and so circumvents most of the hepatic metabolism needed by the other azoles. It is mainly excreted unchanged in urine, although about 10% is metabolized [117]. Fluconazole is a weak substrate for CYP3A4 [10] and a substrate [118] and inhibitor [119] of UGT2B7, both of which have been shown to have genetic variants with altered enzymic activity. However, as only 10% of fluconazole is metabolized, genetic polymorphisms in CYP3A4 or UGT2B7 are unlikely to affect this process significantly.

Caspofungin is not metabolized by phase I enzymes, but is broken down by peptide hydrolysis and N-acetylation. In the plasma, unchanged caspofungin is found mainly in the first 24 hours, but later M0, a linear peptide, is the main metabolite. In urine, in the initial 24-hours, the parent drug is mainly found, but later the hydrolytic metabolites M1 and M2 are the predominant form [120]. N-acetylation is carried out by NATs and numerous genetic polymorphisms affect their activity [113]. However, there is currently no data on which NAT is involved with caspofungin metabolism, and hence whether genetic variability has any impact on it.

Micafungin is metabolized to an M1 metabolite by arylsulfatase, which is then further metabolized by catechol-O-methyltransferase (COMT) to the M2 metabolite [121]. Micafungin's side chain is also metabolized by CYP3A to produce a third metabolite, but this is only a minor metabolic pathway [122]. COMT is present in cells either as a soluble form in the cytoplasm or as a membrane-bound form in the rough endoplasmic reticulum. The soluble form is mainly expressed in peripheral tissues. Both arylsulfatase and COMT are polymorphic, which affects their enzymic activity [113]. Because both of the enzymes involved in micafungin metabolism have genetic variants with differing metabolic activity, the potential exists for this to affect an individual's ability to metabolize micafungin; as yet there are no data as to whether this occurs.

Several antifungal drugs must undergo extensive metabolism before they can be excreted from the body; for them, the potential for genetic variants of the phase I and II enzymes to impact clearance is highest.

Terbinafine is metabolized to five main metabolites by four main metabolic pathways: N-demethylation, deamination, alkyl side chain oxidation, and dihydrodiol formation [45]. The CYP enzymes responsible for each of the metabolic pathways have been defined: N-demethylation is mediated by 2C9, 2C8, and 1A2; deamination, by 3A4; alkyl side chain oxidation, by 1A2, 2C8, 2C9, and 2C19; and dihydrodiol formation by 2C9 and 1A2. Overall, 2C9, 1A2, and 3A4 are the most important isoforms involved in metabolism. All of the CYP isoforms involved are genetically variable [108] and so polymorphisms could affect the metabolism of terbinafine. However, because metabolic activity is spread over a range of isoforms, the potential impact of each polymorphism is limited. As for most other antifungal drugs, there are currently no data on how terbinafine metabolism is affected by CYP polymorphisms.

Of the newer triazoles, posaconazole has the most limited interaction with the CYP enzymes, as it is not metabolized by this route. Indeed, posaconazole does not undergo extensive metabolism, being mainly excreted unchanged, and there are no active metabolites found in plasma [9,32]. The predominant metabolite that is formed is a posaconazole glucuronide produced by UGT1A4 [123]. UGT1A4 is polymorphic, and some of the resultant enzymes have reduced catalytic activity [124]. This, then, is the only point at which genetic polymorphisms might affect posaconazole metabolism, but because most of the drug is excreted unchanged, their potential impact is limited.

Itraconazole is largely metabolized in the liver, with only 3–18% of the parent drug excreted in the feces. The main metabolic pathways include oxidation and N-dealkylation [125], resulting in at least 30 metabolites, of which hydroxy-itraconazole is microbiologically active and present in higher concentrations than the parent drug [19]. Metabolism is mainly carried out by CYP3A4, for which itraconazole is both a substrate and a potent inhibitor [19]. Initial metabolism results in production of hydroxy-itraconazole, which remains attached to the CYP3A4 enzyme complex which further catalyzes the conversion to keto-itraconazole and then N-desalkyl-itraconazole. In addition to itraconazole being a potent CYP3A4 inhibitor, its metabolites have a higher inhibitory capacity than the parent drug, which may explain why such potent CYP3A4 inhibition is seen with itraconazole [126]. Because CYP3A4 is the most abundant isoform in the human intestines, it is estimated that first-pass metabolism may remove up to 50% of the itraconazole dose before it reaches the systemic circulation [18]. Although there are no data in humans, up to 76% of the oral dose and 21% of the IV dose in rats was metabolized by this mechanism, resulting in much lower drug exposure than predicted [127]. This suggests that first-pass metabolism may have significant impact on drug exposure in patients.

Multiple SNPs have been described in CYP3A4, but most of them are low frequency and usually heterozygous with a wild-type allele. However, CYP3A5 is highly polymorphic and, if it is involved in itraconazole metabolism, there is significant potential for genetic variants to cause an impact, both on first-pass and later metabolism; this might go some way to explaining the very variable levels of drug exposure seen in patients. At the moment, there are no data on how these polymorphisms might affect the onward metabolism of hydroxy-itraconazole—the first metabolite formed in the pathway. Because it too is metabolized by CYP3A4, it may be that it has similar potential to be as affected as the parent drug.

In addition, itraconazole is known to be a substrate for UGT1A4, another polymorphic enzyme [118]. The exact role that this enzyme plays in human itraconazole metabolism is not clear, but it too is polymorphic and so is another point at which itraconazole metabolism may be influenced by genetic variability. Although to date there have been no studies examining this area, the data indicating that polymorphisms may be important for itraconazole suggests that such studies may reveal more about itraconazole's pharmacokinetics.

The last of the antifungal drugs to consider is voriconazole, for which there is the most evidence that genetic variability plays a clinically significant role in its metabolism. Voriconazole is extensively metabolized, with over 75–80% excreted via the urine and 20–25% via the feces [22] as multiple metabolites. The main circulating metabolite is voriconazole N-oxide, which is produced by CYP2C19 and CYP3A4, with 2C19 being high-affinity but low-capacity and 3A4 being a low-affinity but high-capacity. CYP2C9 also plays a minor role in voriconazole metabolism [21]. Another metabolite, 4-hydroxyvoriconazole is produced by CYP3A4, and it has been suggested that the abundance of CYP3A4 over CYP2C19 in the liver means that this metabolic pathway is also significant [128]. In addition to CYP-mediated metabolism of voriconazole, which makes up 75% of the metabolic pathway, 25% of voriconazole metabolism is mediated by the FMO enzymes [129]. Although FMOs contribute to the production of voriconazole N-oxide, they are less important than CYP2C19. Voriconazole has recently been found to inhibit UGT2B [130], although the significance of this is unclear.

Patients receiving voriconazole have been shown to have significant variability in their pharmacokinetics, with trough plasma concentrations in one study of 122 patients ranging from <100–9700 mg/L despite the same dosing regimen [131]. CYP2C19 is the most important enzyme involved in voriconazole metabolism and, as discussed previously, its variants give rise to poor metabolizers, heterozygous wild-type metabolizers, and homozygous extensive metabolizers [110]. Voriconazole drug exposure in patients who are poor metabolizers is up to 5 times higher than in those who are homozygous extensive metabolizers [132]. Conversely, the rapid metabolizer phenotype leads to concentrations of voriconazole that are much lower than those seen in extensive or poor metabolizers [133]. Indeed, the variability in drug levels due to polymorphisms in CYP2C19 have been estimated to account for 49% of drug clearance variability. Thus, these polymorphisms are very important in voriconazole pharmacokinetics [134]. However, when multiple CYP genetic polymorphisms coexist, this does not lead to an additive effect on metabolism [135]. In addition to drug exposure variability caused by genetic variants of CYP2C19, there is also variability caused by FMO genetic variants [112], but the extent to which this impacts voriconazole metabolism is unknown.

Because it is known that CYP2C19 variability is important in voriconazole exposure, voriconazole was one of the first drugs for which pharmacogenomic data were included in the manufacturer's FDA submission. It has also been suggested that genome analysis should be performed in the future for all patients receiving voriconazole, in addition to therapeutic drug monitoring to optimize exposure [136].

For voriconazole, more than for any other drug, understanding how genetic polymorphisms affect its metabolism have been pivotal in explaining the pharmacokinetic variability observed in patients who are treated with it.

38.4.4 Elimination

Elimination of drugs from the body is via the liver or kidneys. In the liver drugs are eliminated with bile via the feces, while in the kidneys they are eliminated via urine. Efflux transporters in the liver and kidneys are therefore critical to removal of drugs from the body. Transporters that are responsible for transit of substances across the canalicular membrane and into the bile include P-gp, MRP2, BCRP, BSEP, MDR3, and multidrug and toxin extrusion type transporter 1 (MATE1) [137]. In the kidneys, excretion into urine is mediated via various transporters, including OCTN1, OCTN2, MATE1, MATE2, P-gp, BCRP, MRP2, and MRP4 [138]. Transporters that are responsible for effluxing drugs into the intestinal lumen include P-gp, BCRP, and MRP2 [73], and these may mediate antifungal drug efflux directly into the feces.

Amphotericin B is predominantly eliminated via the bile, with 43% excreted unchanged, while 21% is excreted via the kidneys. Several drug transporters are important in transporting drugs across the canalicular membrane and into the bile, which suggests that amphotericin B must be a substrate of at least one of them. Limited evidence suggests that amphotericin B may be transported via P-gp [139], and this might represent a route for its elimination via both the liver and kidneys. P-gp polymorphisms could potentially impact elimination if indeed amphotericin B is a substrate of P-gp.

Fluconazole and flucytosine are largely excreted unchanged in the urine. Because of their hydrophilic nature and limited protein binding, they may diffuse into the urine passively. However, at the moment there are no data on whether drug transporters contribute to this process, and hence whether genetic variants influence the elimination of either drug.

Elimination of itraconazole is mainly via the feces as metabolites, with some excreted as either metabolites or unchanged in urine. Itraconazole is a substrate of P-gp and CYP3A4, and up to 50% of it may potentially be excreted directly into bile or feces via first-pass metabolism without absorption. Similarly, itraconazole that has been absorbed is likely to be eliminated via P-gp efflux into the intestines. The route of excretion of itraconazole's main metabolite, hydroxy-itraconazole, is not well established and so there are no data as to whether it is eliminated via drug transporters or via passive diffusion. P-gp represents one point at which itraconazole elimination may be affected by genetic polymorphisms.

Voriconazole is mainly excreted as metabolites, with 80% in the urine and about 20% in the feces. It is not a substrate for BCRP, and there is only limited evidence that it is a substrate for P-gp [76,140]. Also, it is currently unknown whether any of voriconazole's metabolites are substrates for drug transporters, but at the moment there is little to suggest that genetic variants impact its elimination.

Posaconazole is predominantly eliminated unchanged in the feces, with some metabolites in the urine. It is a substrate of P-gp, and the presence of large amounts of unchanged drug in the feces suggests that it may undergo some direct efflux from the systemic circulation into the intestines because of P-gp [32]. This may be partly why its absorption is limited. Thus, excretion of posaconazole may be affected by polymorphisms in P-gp, resulting in more or less being left in the circulation depending on the particular effect of the polymorphism.

Caspofungin and anidulafungin are eliminated mainly as metabolites: caspofungin equally in the urine and feces and anidulafungin mainly via the feces. Most of the metabolites for both drugs are peptides, so they are likely to be eliminated by transporters that excrete peptides from the body. It is unknown which transporters eliminate the metabolites from these drugs, and hence whether they have polymorphisms that may affect this process.

Micafungin is eliminated mostly in the bile via the feces. The rate-limiting step in its hepatic elimination is biliary excretion rather than hepatic uptake, and the main drug transporter involved in transport across the canalicular hepatocyte membrane is BSEP, with BCRP having little importance. MRP3 is responsible for transport of micafungin across sinusoidal hepatocyte membranes [100]. Both of these transporters have genetic polymorphisms, and hence there is the potential that the elimination of micafungin is affected by them.

Terbinafine is metabolized to five main metabolites which differ in their excretory route. In plasma, the main metabolite is carboxybutylterbinafine, which, along with the parent drug and three other metabolites, accounts for 80% of the total dose [141]. In urine the major metabolite is demethylcarboxybutylterbinafine, which with the parent drug and three other metabolites accounts for only 14% of the total dose over a 48-hour period. There are currently no data on the interaction of terbinafine with drug transporters, and hence no indication as to whether polymorphisms may affect its elimination.

38.5 CURRENT SUCCESSES IN STRATIFIED MEDICINE

The risk of developing IFD is determined by many factors, including the degree and duration of immunocompromise, the presence of other infections, and underlying disease. As detailed previously, there are various ways of managing patients, including prophylaxis, the enhanced diagnostic approach, and empiric antifungal therapy. Therefore, it has long been understood that not all patients are equally likely to develop infections and so there is a need to quantify each patient's risk. In one study, this concept of risk stratification was formalized in neutropenic patients by dividing them into four categories: low-risk, intermediate low-risk, intermediate high-risk, and high-risk [142]. Patients who had lymphoma, autologous HSCT, or childhood acute lymphoblastic leukemia were considered low-risk; those who had been neutropenic for more than 5 weeks, had graft versus host disease, received steroids at doses >2 mg/kg for more than two weeks, and various others were considered high-risk. Using these, the researchers found that the evidence supported prophylaxis in high-risk groups but not in the intermediate or low-risk groups. This risk stratification was evaluated and found to be valid as a way to define high-risk patients in the clinical setting [143].

Stratification has also been helpful in the interpretation of biomarkers. The *Aspergillus* GM ELISA is used as part of the diagnostic pathway for invasive aspergillosis, but has also been found to have prognostic value. The amount of GM in each sample is expressed as a ratio, with GM indices >0.5 considered positive. The GM index and its rate of decay are also predictive of mortality. If the GM index in the first week increases by 1 unit, the risk of mortality at 6 weeks increases by 25%, whereas each unit of fall decreases the risk of mortality at 6 weeks by 22% [144]. This ability to predict mortality based on the GM index provides another possibility for stratifying patients and in this way appropriately targeting management.

Another developing area in stratified medicine is our understanding of differences in immune recognition mechanisms between individuals that affects their predisposition to fungal disease, particularly invasive aspergillosis. Immune recognition of fungi involves multiple aspects of the immune system. Innate immunity, particularly phagocytic cells, is important in the uptake and clearance of fungi, using a range of ligand–receptor interactions. Triggering of the cytokine release that coordinates the interplay between innate and adaptive immune responses is also important in driving the host response toward either a Th1 proinflammatory and usually protective response, or toward a Th2 anti-inflammatory and often unfavorable response. Polymorphisms within many of the genes coding for the receptors and cytokines have been described and found to correlate with changes in IFD risk. For example, one study found that the presence of two specific haplotypes for TLR4 increased the risk of invasive aspergillosis in HSCT recipients [145]. Our understanding of this area is expanding rapidly, and with it our ability to define which patients are likely to develop IFD and so tailor their care to minimize the risk of disease.

38.6 FUTURE PERSPECTIVES

Successful treatment of IFDs relies on patients receiving an antifungal that achieves sufficient drug exposure at the site of infection; therefore, improving the absorption and distribution of drugs is of paramount importance. As we understand more about the pharmacokinetics of antifungals, it is becoming clear that many interact with genetically polymorphic enzymes and/or transporters, which means that there is significant potential for these polymorphisms to affect a drug's journey through the body. For some drugs, such as voriconazole, polymorphisms in the CYP2C19 enzyme cause most of the variability in drug exposure between patients. For other drugs, there is less evidence for this, but this is probably a result of the lack of relevant studies. A priority now should be to systematically examine the pharmacogenomic data that do exist for all antifungal drugs and conduct clinical studies, including appropriately chosen populations with relevant polymorphisms, to determine how much these polymorphisms impact clinical patient outcomes. Understanding this is pivotal to our ability to tailor the management of each patient and provide truly personalized medicine.

Once such studies have been conducted, the use of next-generation sequencing [146] will allow screening for polymorphisms as part of a patient's diagnostic workup. For some patients who undergo intensive diagnostic investigations before procedures—for example, tissue typing before HSCT or solid organ transplantation—the addition of screening for specific polymorphisms involved in a patient's susceptibility to either IFD or drug pharmacokinetics is an obvious additional step. The ability to use the right drug at the right dose, limiting toxicity but achieving efficacy—all informed by an understanding of a patient's genetic profile—is becoming a more achievable goal. Ultimately this will not only provide personalized medicine for the patient, but also save resources in an era in which financial constraints increasingly influence patient care.

REFERENCES

[1] Oxford AE, Raistrick H, Simonart P. Studies in the biochemistry of micro-organisms: Griseofulvin, C(17)H(17)O(6)Cl, a metabolic product of Penicillium griseo-fulvum Dierckx. Biochem J 1939;33(2):240–8.

[2] Gallis HA, Drew RH, Pickard WW. Amphotericin B: 30 years of clinical experience. Rev Infect Dis 1990;12(2):308–29.

[3] Bellmann R. Clinical pharmacokinetics of systemically administered antimycotics. Curr Clin Pharmacol 2007;2(1):37–58.

[4] Laniado-Laborin R, Cabrales-Vargas MN. Amphotericin B: side effects and toxicity. Rev Iberoam Micol 2009;26(4):223–7.

[5] Herbrecht R, Natarajan-Ame S, Nivoix Y, Letscher-Bru V. The lipid formulations of amphotericin B. Expert Opin Pharmacother 2003;4(8):1277–87.

[6] Bekersky I, Fielding RM, Dressler DE, Lee JW, Buell DN, Walsh TJ. Pharmacokinetics, excretion, and mass balance of liposomal amphotericin B (AmBisome) and amphotericin B deoxycholate in humans. Antimicrob Agents Chemother 2002;46(3):828–33.

[7] Ostrosky-Zeichner L, Marr KA, Rex JH, Cohen SH. Amphotericin B: time for a new "gold standard". Clin Infect Dis 2003;37:415–25.

[8] Woolley DW. Some biological effects produced by benzimidazole and their reversal by purines. J Biol Chem 1944;152:225–32.

[9] Dodds Ashley ES. Pharmacology of azole antifungal agents. In: Antifungal therapy. Informa Healthcare; 2010. pp. 199–218.

[10] Venkatakrishnan K, von Moltke LL, Greenblatt DJ. Effects of the antifungal agents on oxidative drug metabolism: clinical relevance. Clin Pharmacokinet 2000;38(2):111–80.

[11] Sakaeda T, Iwaki K, Kakumoto M, Nishikawa M, Niwa T, Jin JS. Effect of micafungin on cytochrome P450 3A4 and multidrug resistance protein 1 activities, and its comparison with azole antifungal drugs. J Pharm Pharmacol 2005;57(6):759–64.

[12] Niwa T, Shiraga T, Takagi A. Effect of antifungal drugs on cytochrome P450 (CYP) 2C9, CYP2C19, and CYP3A4 activities in human liver microsomes. Biol Pharm Bull 2005;28(9):1805–8.

[13] Brammer KW, Farrow PR, Faulkner JK. Pharmacokinetics and tissue penetration of fluconazole in humans. Rev Infect Dis 1990;12(Suppl.):318–27.

[14] Wang EJ, Lew K, Casciano CN, Clement RP, Johnson WW. Interaction of common azole antifungals with P glycoprotein. Antimicrob Agents Chemother 2002;46(1):160–5.

[15] Bruggemann RJM, Alffenaar JWC, Blijlevens NMA, Billaud EM, Kosterink JGW, Verweij PE. Clinical relevance of the pharmacokinetic interactions of azole antifungal drugs with other coadministered agents. Clin Infect Dis 2009;48(10):1441–58.

[16] Sabatelli F, Patel R, Mann PA, Mendrick CA, Norris CC, Hare R. *In vitro* activities of posaconazole, fluconazole, itraconazole, voriconazole, and amphotericin B against a large collection of clinically important molds and yeasts. Antimicrob Agents Chemother 2006;50(6):2009–15.

[17] Espinel-Ingroff A. *In vitro* fungicidal activities of voriconazole, itraconazole, and amphotericin B against opportunistic moniliaceous and dematiaceous fungi. J Clin Microbiol 2001;39(3):954–8.

[18] Lestner J, Hope WW. Itraconazole: an update on pharmacology and clinical use for treatment of invasive and allergic fungal infections. Expert Opin Drug Metab Toxicol 2013;9(7):911–26.

[19] Templeton IE, Thummel KE, Kharasch ED, Kunze KL, Hoffer C, Nelson WL. Contribution of itraconazole metabolites to inhibition of CYP3A4 *in vivo*. Clin Pharmacol Ther 2008;83(1):77–85.

[20] Purkins L, Wood N, Greenhalgh K, Allen MJ, Oliver SD. Voriconazole, a novel wide-spectrum triazole: oral pharmacokinetics and safety. Br J Clin Pharmacol 2003;56(Suppl. 1):10–16.

[21] Hyland R, Jones BC, Smith DA. Identification of the cytochrome P450 enzymes involved in the N-oxidation of voriconazole. Drug Metab Dispos 2003;31(5):540–7.

[22] Roffey SJ, Cole S, Comby P, Gibson D, Jezequel SG, Nedderman AN. The disposition of voriconazole in mouse, rat, rabbit, guinea pig, dog, and human. Drug Metab Dispos 2003;31(6):731–41.

[23] Pasqualotto AC, Xavier MO, Andreolla HF, Linden R. Voriconazole therapeutic drug monitoring: focus on safety. Expert Opin Drug Saf 2010;9(1):125–37.

[24] Vadnerkar A, Nguyen MH, Mitsani D, Crespo M, Pilewski J, Toyoda Y. Voriconazole exposure and geographic location are independent risk factors for squamous cell carcinoma of the skin among lung transplant recipients. J Heart Lung Transplant 2010;29(11):1240–4.

[25] Tan K, Brayshaw N, Tomaszewski K, Troke P, Wood N. Investigation of the potential relationships between plasma voriconazole concentrations and visual adverse events or liver function test abnormalities. J Clin Pharmacol 2006;46(2):235–43.

[26] Jeong S, Nguyen PD, Desta Z. Comprehensive *in vitro* analysis of voriconazole inhibition of eight cytochrome P450 (CYP) enzymes: major effect on CYPs 2B6, 2C9, 2C19, and 3A. Antimicrob Agents Chemother 2009;53(2):541–51.

[27] Krishna G, Ma L, Martinho M, O'Mara E. A single dose phase I study to evaluate the pharmacokinetics of posaconazole new tablet and capsule formulations relative to oral suspension. Antimicrob Agents Chemother 2012;56(8):4196–201.

[28] Soczo G, Kardos G, McNicholas PM, Balogh E, Gergely L, Varga I. Correlation of posaconazole minimum fungicidal concentration and time kill test against nine Candida species. J Antimicrob Chemother 2007;60(5):1004–9.

[29] Krishna G, Moton A, Ma L, Medlock MM, McLeod J. Pharmacokinetics and absorption of posaconazole oral suspension under various gastric conditions in healthy volunteers. Antimicrob Agents Chemother 2009;53(3):958–66.

[30] Dolton MJ, Ray JE, Chen SC, Ng K, Pont L, McLachlan AJ. Multicenter study of posaconazole therapeutic drug monitoring: exposure-response relationship and factors affecting concentration. Antimicrob Agents Chemother 2012;56(11):5503–10.

[31] Wexler D, Courtney R, Richards W, Banfield C, Lim J, Laughlin M. Effect of posaconazole on cytochrome P450 enzymes: a randomized, open-label, two-way crossover study. Eur J Pharm Sci 2004;21(5):645–53.

[32] Krieter P, Flannery B, Musick T, Gohdes M, Martinho M, Courtney R. Disposition of posaconazole following single-dose oral administration in healthy subjects. Antimicrob Agents Chemother 2004;48(9):3543–51.

[33] Johnson MD, Mohr J. Echinocandins for prevention and treatment of invasive fungal infections. In: Antifungal therapy. Informa Healthcare; 2010. pp. 219–42.

[34] Wagner C, Graninger W, Presterl E, Joukhadar C. The echinocandins: comparison of their pharmacokinetics, pharmacodynamics and clinical applications. Pharmacology 2006;78(4):161–77.

[35] Denning DW. Echinocandin antifungal drugs. Lancet 2003;363:1142–51.

[36] Colburn DE, Giles FJ, Oladovich D, Smith JA. *In vitro* evaluation of cytochrome P450-mediated drug interactions between cytarabine, idarubicin, itraconazole and caspofungin. Hematology 2004;9(3):217–21.

[37] Vermes A, Guchelaar HJ, Dankert J. Flucytosine: a review of its pharmacology, clinical indications, pharmacokinetics, toxicity and drug interactions. J Antimicrob Chemother 2000;46:171–9.

[38] Drew RH. Flucytosine. In: Antifungal therapy. Informa Healthcare; 2010. pp. 184–98.

[39] Vermes A, Guchelaar HJ, van Kuilenberg ABP, Dankert J. 5-fluorocytosine-related bone-marrow depression and conversion to fluorouracil: a pilot study. Fundam Clin Pharmacol 2002;16(1):39–47.

[40] Whelan WL. The genetic basis of resistance to 5-fluorocytosine in Candida species and Cryptococcus neoformans. Crit Rev Microbiol 1987;15(1):45–56.

[41] Jain S, Sehgal VN. Terbinafine, a unique oral antifungal: current perceptions. Int J Dermatol 2000;39(6):412–23.

[42] Schuster I, Ryder NS. Allylamines – mode and selectivity of action compared to azole antifungals and biological fate in mammalian organisms. J Dermatol Treat 1990;1(Suppl. 2):7–9.

[43] Ryder NS, Leitner I. *In vitro* activity of terbinafine (Lamisil): an update. J Dermatol Treat 1998;9(Suppl. 1):S23–8.

[44] Jensen JC. Clinical pharmacokinetics of terbinafine (Lamisil). Clin Exp Dermatol 1989;14(2):110–3.

[45] Vickers AE, Sinclair JR, Zollinger M, Heitz F, Glanzel U, Johanson L. Multiple cytochrome P-450s involved in the metabolism of terbinafine suggest a limited potential for drug-drug interactions. Drug Metab Dispos 1999;27(9):1029–38.

[46] Suhonen R, Neuvonen PJ. The tolerability profile of terbinafine. Rev Contemp Pharmacother 1997;8:373–86.

[47] Ahonen J, Olkkola KT, Neuvonen PJ. Effect of itraconazole and terbinafine on the pharmacokinetics and pharmacodynamics of midazolam in healthy volunteers. Br J Clin Pharmacol 1995;40(3):270–2.

[48] Abdel-Rahman SM, Marcucci K, Boge T, Gotschall RR, Kearns GL, Leeder JS. Potent inhibition of cytochrome P-450 2D6-mediated dextromethorphan O-demethylation by terbinafine. Drug Metab Dispos 1999;27(7):770–5.

[49] Hynninen VV, Olkkola KT, Bertilsson L, Kurkinen K, Neuvonen PJ, Laine K. Effect of terbinafine and voriconazole on the pharmacokinetics of the antidepressant venlafaxine. Clin Pharmacol Ther 2008;83(2):342–8.

[50] Castberg I, Helle J, Aamo TO. Prolonged pharmacokinetic drug interaction between terbinafine and amitriptyline. Ther Drug Monit 2005;27(5):680–2.

[51] Burzykowski T, Molenberghs G, Abeck D, Haneke E, Hay RJ, Katsambas A. High prevalence of foot diseases in Europe; results of the Achilles Project. Mycoses 2003;46:496–505.

[52] Fuller LC, Child FJ, Midgley G, Higgins EM. Diagnosis and management of scalp ringworm. BMJ 2003;326(7388):539–41.

[53] Lopez-Martinez R. Candidosis, a new challenge. Clin Dermatol 2010;28(2):178–84.

[54] Richardson MD, Lass-Florl C. Changing epidemiology of systemic fungal infections. Clin Microbiol Infect 2008;14(Supp. 4):5–24.

[55] Menzin J, Meyers JL, Friedman M, Perfect JR, Langston AA, Danna RP. Mortality, length of hospitalization, and costs associated with invasive fungal infections in high-risk patients. Am J Health Syst Pharm 2009;66(19):1711–7.

[56] Verweij PE, Snelders E, Kema GH, Mellado E, Melchers WJ. Azole resistance in Aspergillus fumigatus: a side-effect of environmental fungicide use? Lancet Infect Dis 2009;9(12):789–95.

[57] Howard SJ, Webster I, Moore CB, Gardiner RE, Park S, Perlin DS. Multi-azole resistance in Aspergillus fumigatus. Int J Antimicrob Agents 2006;28(5):450–3.

[58] Pfaller MA, Diekema DJ. Epidemiology of invasive mycoses in North America. Crit Rev Microbiol 2010;36(1):1–53.

[59] Ashbee HR, Gilleece MH. Has the era of individualised medicine arrived for antifungals? A review of antifungal pharmacogenomics. Bone Marrow Transplant 2011;25(47):881–94.

[60] Marr KA, Laverdiere M, Gugel A, Leisenring W. Antifungal therapy decreases sensitivity of the Aspergillus galactomannan enzyme immunoassay. Clin Infect Dis 2005;40(12):1762–9.

[61] Barnes RA, White PL, Bygrave C, Evans N, Healy B, Kell J. Clinical impact of enhanced diagnosis of invasive fungal disease in high-risk haematology and stem cell transplant patients. J Clin Pathol 2009;62(1):64–9.

[62] Cuenca-Estrella M, Bernal-Martinez L, Buitrago MJ, Castelli MV, Gomez-Lopez A, Zaragoza O. Update on the epidemiology and diagnosis of invasive fungal infection. Int J Antimicrob Agents 2008;32(Suppl. 2):S143–7.

[63] Garey KW, Rege M, Pai MP, Mingo DE, Suda KJ, Turpin RS. Time to initiation of fluconazole therapy impacts mortality in patients with candidemia: a multi-institutional study. Clin Infect Dis 2006;43(1):25–31.

[64] Meletiadis J, Chanock S, Walsh TJ. Human pharmacogenomic variations and their implications for antifungal efficacy. Clin Microbiol Rev 2006;19(4):763–87.

[65] Johnson MD. Pharmacogenomics of systemic antifungal agents. Curr Fungal Infect Rep 2009;3:111–6.

[66] Prentice AG, Glasmacher A. Making sense of itraconazole pharmacokinetics. J Antimicrob Chemother 2005;56(Suppl. 1):i17–22.

[67] Courtney R, Radwanski E, Lim J, Laughlin M. Pharmacokinetics of posaconazole coadministered with antacid in fasting or nonfasting healthy men. Antimicrob Agents Chemother 2004;48(3):804–8.

[68] Williams MP, Sercombe J, Hamilton MI, Pounder RE. A placebo-controlled trial to assess the effects of 8 days of dosing with rabeprazole versus omeprazole on 24-h intragastric acidity and plasma gastrin concentrations in young healthy male subjects. Aliment Pharmacol Ther 1998;12(11):1079–89.

[69] Feldman M, Barnett C. Fasting gastric pH and its relationship to true hypochlorhydria in humans. Dig Dis Sci 1991;36(7):866–9.

[70] Welage LS, Carver PL, Revankar S, Pierson C, Kauffman CA. Alterations in gastric acidity in patients infected with human immunodeficiency virus. Clin Infect Dis 1995;21(6):1431–8.

[71] Brett S. Science review: the use of proton pump inhibitors for gastric acid suppression in critical illness. Crit Care 2005;9(1):45–50.

[72] Grube RR, May DB. Stress ulcer prophylaxis in hospitalized patients not in intensive care units. Am J Health Syst Pharm 2007;64(13):1396–400.

[73] Estudante M, Morais JG, Soveral G, Benet LZ. Intestinal drug transporters: an overview. Adv Drug Deliv Rev 2013;65(10):1340–56.

[74] Tucker TG, Milne AM, Fournel-Gigleux S, Fenner KS, Coughtrie MW. Absolute immunoquantification of the expression of ABC transporters P-glycoprotein, breast cancer resistance protein and multidrug resistance-associated protein 2 in human liver and duodenum. Biochem Pharmacol 2012;83(2):279–85.

[75] Schwab M, Eichelbaum M, Fromm MF. Genetic polymorphisms of the human MDR1 drug transporter. Annu Rev Pharmacol Toxicol 2003;43:285–307.

[76] Wakiec R, Prasad R, Morschhauser J, Barchiesi F, Borowski E, Milewski S. Voriconazole and multidrug resistance in Candida albicans. Mycoses 2007;50(2):109–15.

[77] Gupta A, Unadkat JD, Mao Q. Interactions of azole antifungal agents with the human breast cancer resistance protein (BCRP). J Pharm Sci 2007;96(12):3226–35.

[78] Courtney R, Sansone-Parsons A, Devlin D, Soni P, Laughlin M, Simon J. P-glycoprotein (P-gp) expression and genotype: exploratory analysis of posaconazole (POS) in healthy volunteers. ICAAC Abstract A-40; 2004.

[79] Arredondo G, Calvo R, Marcos F, Martinez-Jorda R, Suarez E. Protein binding of itraconazole and fluconazole in patients with cancer. Int J Clin Pharmacol Ther 1995;33(8):449–52.

[80] Brajtburg J, Elberg S, Bolard J, Kobayashi GS, Levy RA, Ostlund Jr RE. Interaction of plasma proteins and lipoproteins with amphotericin B. J Infect Dis 1984;149(6):986–97.

[81] Bekersky I, Fielding RM, Dressler DE, Lee JW, Buell DN, Walsh TJ. Plasma protein binding of amphotericin B and pharmacokinetics of bound versus unbound amphotericin B after administration of intravenous liposomal amphotericin B (AmBisome) and amphotericin B deoxycholate. Antimicrob Agents Chemother 2002;46(3):834–40.

[82] Kragh-Hansen U, Minchiotti L, Galliano M, Peters Jr T. Human serum albumin isoforms: genetic and molecular aspects and functional consequences. Biochim Biophys Acta 2013;1830(8):5405–17.

[83] Nishi K, Ueno M, Murakami Y, Fukunaga N, Akuta T, Kadowaki D. A site-directed mutagenesis study of drug-binding selectivity in genetic variants of human alpha(1)-acid glycoprotein. J Pharm Sci 2009;98(11):4316–26.

[84] Ho RH, Kim RB. Transporters and drug therapy: implications for drug disposition and disease. Clin Pharmacol Ther 2005;78(3):260–77.

[85] Valeriote F, Medoff G, Dieckman J. Potentiation of anticancer agent cytotoxicity against sensitive and resistant AKR leukemia by amphotericin B1. Cancer Res 1979;39(6 Pt 1):2041–5.

[86] Ishizaki J, Ito S, Jin M, Shimada T, Ishigaki T, Harasawa Y. Mechanism of decrease of oral bioavailability of cyclosporin A during immunotherapy upon coadministration of amphotericin B. Biopharm Drug Dispos 2008;29(4):195–203.

[87] Miyama T, Takanaga H, Matsuo H, Yamano K, Yamamoto K, Iga T. P-glycoprotein-mediated transport of itraconazole across the blood-brain barrier. Antimicrob Agents Chemother 1998;42(7):1738–44.

[88] Imbert F, Jardin M, Fernandez C, Gantier JC, Dromer F, Baron G. Effect of efflux inhibition on brain uptake of itraconazole in mice infected with Cryptococcus neoformans. Drug Metab Dispos 2003;31(3):319–25.

[89] Yoshikado T, Takada T, Yamamoto T, Yamaji H, Ito K, Santa T. Itraconazole-induced cholestasis: involvement of the inhibition of bile canalicular phospholipid translocator MDR3/ABCB4. Mol Pharmacol 2011;79(2):241–50.

[90] Lennernas H. Clinical pharmacokinetics of atorvastatin. Clin Pharmacokinet 2003;42(13):1141–60.

[91] Uno T, Shimizu M, Sugawara K, Tateishi T. Lack of dose-dependent effects of itraconazole on the pharmacokinetic interaction with fexofenadine. Drug Metab Dispos 2006;34(11):1875–9.

[92] Hirano M, Maeda K, Shitara Y, Sugiyama Y. Drug-drug interaction between pitavastatin and various drugs via OATP1B1. Drug Metab Dispos 2006;34(7):1229–36.

[93] Yamaguchi H, Takeuchi T, Okada M, Kobayashi M, Unno M, Abe T. Screening of antibiotics that interact with organic anion-transporting polypeptides 1B1 and 1B3 using fluorescent probes. Biol Pharm Bull 2011;34(3):389–95.

[94] Lutsar I, Roffey S, Troke P. Voriconazole concentrations in the cerebrospinal fluid and brain tissue of guinea pigs and immunocompromised patients. Clin Infect Dis 2003;37(5):728–32.

[95] Weiler S, Fiegl D, MacFarland R, Stienecke E, Bellmann-Weiler R, Dunzendorfer S. Human tissue distribution of voriconazole. Antimicrob Agents Chemother 2011;55(2):925–8.

[96] Li Y, Theuretzbacher U, Clancy CJ, Nguyen MH, Derendorf H. Pharmacokinetic/pharmacodynamic profile of posaconazole. Clin Pharmacokinet 2010;49(6):379–96.

[97] Stone JA, Xu X, Winchell GA, Deutsch PJ, Pearson PG, Migoya EM. Disposition of caspofungin: role of distribution in determining pharmacokinetics in plasma. Antimicrob Agents Chemother 2004;48(3):815–23.

[98] Sandhu P, Lee W, Xu X, Leake BF, Yamazaki M, Stone JA. Hepatic uptake of the novel antifungal agent caspofungin. Drug Metab Dispos 2005;33(5):676–82.

[99] Stone JA, Migoya EM, Hickey L, Winchell GA, Deutsch PJ, Ghosh K. Potential for interactions between caspofungin and nelfinavir or rifampin. Antimicrob Agents Chemother 2004;48(11):4306–14.

[100] Yanni SB, Augustijns PF, Benjamin Jr DK, Brouwer KL, Thakker DR, Annaert PP. *In vitro* investigation of the hepatobiliary disposition mechanisms of the antifungal agent micafungin in humans and rats. Drug Metab Dispos 2010;38(10):1848–56.

[101] Ho RH, Leake BF, Roberts RL, Lee W, Kim RB. Ethnicity-dependent polymorphism in Na+-taurocholate cotransporting polypeptide (SLC10A1) reveals a domain critical for bile acid substrate recognition. J Biol Chem 2004;279(8):7213–22.

[102] Inskeep PB, Walsky RL, Feng B, Campbell S. Lack of anidulafungin interactions with CYP enzymes and transporters in *in vivo* and *in vitro* systems. In: Proceedings of the 18th European congress of clinical microbiology and infectious disease, Abst P1049; 2008.

[103] Badyal DK, Dadhich AP. Cytochrome P450 and drug interactions. Indian J Pharmacol 2001;33:248–59.

[104] Paine MF, Hart HL, Ludington SS, Haining RL, Rettie AE, Zeldin DC. The human intestinal cytochrome P450 “pie”. Drug Metab Dispos 2006;34(5):880–6.

[105] Shimada T, Yamazaki H, Mimura M, Inui Y, Guengerich FP. Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. J Pharmacol Exp Ther 1994;270(1):414–23.

[106] Paine MF, Khalighi M, Fisher JM, Shen DD, Kunze KL, Marsh CL. Characterization of interintestinal and intraintestinal variations in human CYP3A-dependent metabolism. J Pharmacol Exp Ther 1997;283(3):1552–62.

[107] Hsieh KP, Lin YY, Cheng CL, Lai ML, Lin MS, Siest JP. Novel mutations of CYP3A4 in Chinese. Drug Metab Dispos 2001;29(3):268–73.

[108] <www.cypalleles ki se/> [last accessed 20.08.13].

[109] Stubbins MJ, Harries LW, Smith G, Tarbit MH, Wolf CR. Genetic analysis of the human cytochrome P450 CYP2C9 locus. Pharmacogenetics 1996;6(5):429–39.

[110] Desta Z, Zhao X, Shin JG, Flockhart DA. Clinical significance of the cytochrome P450 2C19 genetic polymorphism. Clin Pharmacokinet 2002;41(12):913–58.

[111] Lamba JK, Lin YS, Schuetz EG, Thummel KE. Genetic contribution to variable human CYP3A-mediated metabolism. Adv Drug Deliv Rev 2002;54(10):1271–94.

[112] Cashman JR, Zhang J. Human flavin-containing monooxygenases. Annu Rev Pharmacol Toxicol 2006;46:65–100.

[113] Jancova P, Anzenbacher P, Anzenbacherova E. Phase II drug metabolizing enzymes. Biomed Pap 2010;154(2):103–16. [Palacký University, Faculty of Medicine and Dentistry, Olomouc, Czech Republic].

[114] Damle BD, Dowell JA, Walsky RL, Weber GL, Stogniew M, Inskeep PB. *In vitro* and *in vivo* studies to characterize the clearance mechanism and potential cytochrome P450 interactions of anidulafungin. Antimicrob Agents Chemother 2009;53(3):1149–56.

[115] Polak A, Eschenhof E, Fernex M, Scholer HJ. Metabolic studies with 5-fluorocytosine-6-14C in mouse, rat, rabbit, dog and man. Chemotherapy 1976;22(3–4):137–53.

[116] Brockmeyer NH, Gambichler T, Bader A, Kreuter A, Kurowski M, Sander P. Impact of amphotericin B on the cytochrome P450 system in HIV-infected patients. Eur J Med Res 2004;9(2):51–4.

[117] Humphrey MJ, Jevons S, Tarbit MH. Pharmacokinetic evaluation of UK-49,858, a metabolically stable triazole antifungal drug, in animals and humans. Antimicrob Agents Chemother 1985;28(5):648–53.

[118] Bourcier K, Hyland R, Kempshall S, Jones R, Maximilien J, Irvine N. Investigation into UDP-glucuronosyltransferase (UGT) enzyme kinetics of imidazole- and triazole-containing antifungal drugs in human liver microsomes and recombinant UGT enzymes. Drug Metab Dispos 2010;38(6):923–9.

[119] Uchaipichat V, Winner LK, Mackenzie PI, Elliot DJ, Williams JA, Miners JO. Quantitative prediction of *in vivo* inhibitory interactions involving glucuronidated drugs from *in vitro* data: the effect of fluconazole on zidovudine glucuronidation. Br J Clin Pharmacol 2006;61(4):427–39.

[120] Balani SK, Xu X, Arison BH, Silva MV, Gries A, DeLuna FA. Metabolites of caspofungin acetate, a potent antifungal agent, in human plasma and urine. Drug Metab Dispos 2000;28(11):1274–8.

[121] Hebert MF, Smith HE, Marbury TC, Swan SK, Smith WB, Townsend RW. Pharmacokinetics of micafungin in healthy volunteers, volunteers with moderate liver disease, and volunteers with renal dysfunction. J Clin Pharmacol 2005;45(10):1145–52.

[122] Kofla G, Ruhnke M. Pharmacology and metabolism of anidulafungin, caspofungin and micafungin in the treatment of invasive candidosis: review of the literature. Eur J Med Res 2011;16(4):159–66.

[123] Ghosal A, Hapangama N, Yuan Y, Achanfuo-Yeboah J, Iannucci R, Chowdhury S. Identification of human UDP-glucuronosyltransferase enzyme(s) responsible for the glucuronidation of posaconazole (Noxafil). Drug Metab Dispos 2004;32(2):267–71.

[124] Mori A, Maruo Y, Iwai M, Sato H, Takeuchi Y. UDP-glucuronosyltransferase 1A4 polymorphisms in a Japanese population and kinetics of clozapine glucuronidation. Drug Metab Dispos 2005;33(5):672–5.

[125] Heykants J, van Peer A, Van de Velde V, Van Rooy P, Meuldermans W, Lavrijsen K. The clinical pharmacokinetics of itraconazole: an overview. Mycoses 1989;32(Supp. 1):67–87.

[126] Isoherranen N, Kunze KL, Allen KE, Nelson WL, Thummel KE. Role of itraconazole metabolites in CYP3A4 inhibition. Drug Metab Dispos 2004;32(10):1121–31.

[127] Yoo SD, Kang E, Jun H, Shin BS, Lee KC, Lee KH. Absorption, first-pass metabolism, and disposition of itraconazole in rats. Chem Pharm Bull (Tokyo) 2000;48(6):798–801.

[128] Murayama N, Imai N, Nakane T, Shimizu M, Yamazaki H. Roles of CYP3A4 and CYP2C19 in methyl hydroxylated and N-oxidized metabolite formation from voriconazole, a new anti-fungal agent, in human liver microsomes. Biochem Pharmacol 2007;73(12):2020–6.

[129] Yanni SB, Annaert PP, Augustijns P, Bridges A, Gao Y, Benjamin Jr DK. Role of flavin-containing monooxygenase in oxidative metabolism of voriconazole by human liver microsomes. Drug Metab Dispos 2008;36(6):1119–25.

[130] Frechen S, Junge L, Saari TI, Suleiman AA, Rokitta D, Neuvonen PJ. A semiphysiological population pharmacokinetic model for dynamic inhibition of liver and gut wall cytochrome P450 3A by voriconazole. Clin Pharmacokinet 2013;52(10):763–81.

[131] Denning DW, Ribaud P, Milpied N, Caillot D, Herbrecht R, Thiel E. Efficacy and safety of voriconazole in the treatment of acute invasive aspergillosis. Clin Infect Dis 2002;34(5):563–71.

[132] Ikeda Y, Umemura K, Kondo K, Sekiguchi K, Miyoshi S, Nakashima M. Pharmacokinetics of voriconazole and cytochrome P450 2C19 genetic status. Clin Pharmacol Ther 2004;75(6):587–8.

[133] Wang G, Lei HP, Li Z, Tan ZR, Guo D, Fan L. The CYP2C19 ultra-rapid metabolizer genotype influences the pharmacokinetics of voriconazole in healthy male volunteers. Eur J Clin Pharmacol 2009;65(3):281–5.

[134] Weiss J, Ten Hoevel MM, Burhenne J, Walter-Sack I, Hoffmann MM, Rengelshausen J. CYP2C19 genotype is a major factor contributing to the highly variable pharmacokinetics of voriconazole. J Clin Pharmacol 2009;49(2):196–204.

[135] Geist MJ, Egerer G, Burhenne J, Mikus G. Safety of voriconazole in a patient with CYP2C9*2/CYP2C9*2 genotype. Antimicrob Agents Chemother 2006;50(9):3227–8.

[136] Mikus G, Scholz IM, Weiss J. Pharmacogenomics of the triazole antifungal agent voriconazole. Pharmacogenomics 2011;12(6):861–72.

[137] Degorter MK, Kim RB. Hepatic drug transporters, old and new: pharmacogenomics, drug response, and clinical relevance. Hepatology 2009;50(4):1014–6.

[138] Morrissey KM, Stocker SL, Wittwer MB, Xu L, Giacomini KM. Renal transporters in drug development. Annu Rev Pharmacol Toxicol 2013;53:503–29.

[139] Risovic V, Sachs-Barrable K, Boyd M, Wasan KM. Potential mechanisms by which Peceol increases the gastrointestinal absorption of amphotericin B. Drug Dev Ind Pharm 2004;30(7):767–74.

[140] Nivoix Y, Leveque D, Herbrecht R, Koffel JC, Beretz L, Ubeaud-Sequier G. The enzymatic basis of drug-drug interactions with systemic triazole antifungals. Clin Pharmacokinet 2008;47(12):779–92.

[141] Humbert H, Cabiac MD, Denouel J, Kirkesseli S. Pharmacokinetics of terbinafine and of its five main metabolites in plasma and urine, following a single oral dose in healthy subjects. Biopharm Drug Dispos 1995;16(8):685–94.

[142] Prentice HG, Kibbler CC, Prentice AG. Towards a targeted, risk-based, antifungal strategy in neutropenic patients. Br J Haematol 2000;110(2):273–84.

[143] McLintock LA, Jordanides NE, Allan EK, Copland M, Stewart K, Parker A. The use of a risk group stratification in the management of invasive fungal infection: a prospective evaluation. Br J Haematol 2004;124(3):403–4.

[144] Koo S, Bryar JM, Baden LR, Marty FM. Prognostic features of galactomannan antigenemia in galactomannan-positive invasive aspergillosis. J Clin Microbiol 2010;48(4):1255–60.

[145] Ok M, Einsele H, Loeffler J. Genetic susceptibility to Aspergillus fumigatus infections. Int J Med Microbiol 2011;301(5):445–52.

[146] Miller FA, Hayeems RZ, Bytautas JP, Bedard PL, Ernst S, Hirte H. Testing personalized medicine: patient and physician expectations of next-generation genomic sequencing in late-stage cancer care. Eur J Hum Genet 2014;22(3):391–5.

Chapter 39

Advances in Understanding and Treatment of Human African Trypanosomiasis: Divergent Diseases Caused by Distinct Parasites

Federica Giordani, Roy Mwenechanya, and Michael P. Barrett
Wellcome Trust Centre for Molecular Parasitology and Glasgow Polyomics,
Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary
and Life Sciences, University of Glasgow, UK

Chapter Outline

39.1 Overview 901
39.2 The Genetics of Trypanosome Lytic Factors and Parasite Factors Enabling Protection Against Lysis 903
39.3 Immune Responses and Immunoavoidance 905
39.4 Pathologies Associated with Hat 906
39.5 Diagnosis of Hat 907
39.6 Staging the Disease 908
39.7 Treatment 909
  39.7.1 Current Drugs 909
  39.7.2 Stage 1 910
  39.7.3 Stage 2 911
  39.7.4 Compounds in Clinical or Advanced Preclinical Trials 913
39.8 Conclusions and Future Perspectives 913
Glossary 914
References 914

39.1 OVERVIEW

Human African trypanosomiasis (HAT), also commonly known as sleeping sickness in its second, CNS-involved stage, is endemic to Sub-Saharan Africa (Figure 39.1) but also reported, although very rarely, beyond the African subcontinent [1]. The disease has been known to mankind for many years [2] and most probably played a key role in shaping early evolution, given that the causative trypanosome parasites are found in those regions of Sub-Saharan Africa where human beings are proposed to have first evolved [3]. Two subspecies of *Trypanosoma brucei* are commonly pathogenic for humans: *T. b. gambiense* and *T. b. rhodesiense*. The two diseases present different epidemiology, the causative parasites can be distinguished at the molecular level, and the clinical progression of disease caused by both is distinct [1,2]. Furthermore, treatment of rhodesiense or gambiense disease is different, and separate diagnostic algorithms are required for both [4]. It has recently become clear that it is possible for other African trypanosome subspecies to cause disease in humans, but only in rare cases where individuals are deficient in trypanosome lytic factor(s) that kill the trypanosome species that have not evolved specific mechanisms to resist lysis [5].

T. b. rhodesiense causes acute and rapidly progressive infection in Eastern and Southern Africa. *T. b. gambiense* infection is found in West and Central Africa and progresses more slowly. Although the two subspecies do not colocalize geographically at the moment, *T. b. rhodesiense* has been spreading northward in Uganda and there is a risk that it will reach areas currently occupied by *T. b. gambiense* [6]. A large majority of HAT infections are today due to

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00039-6
© 2014 Elsevier Inc. All rights reserved.

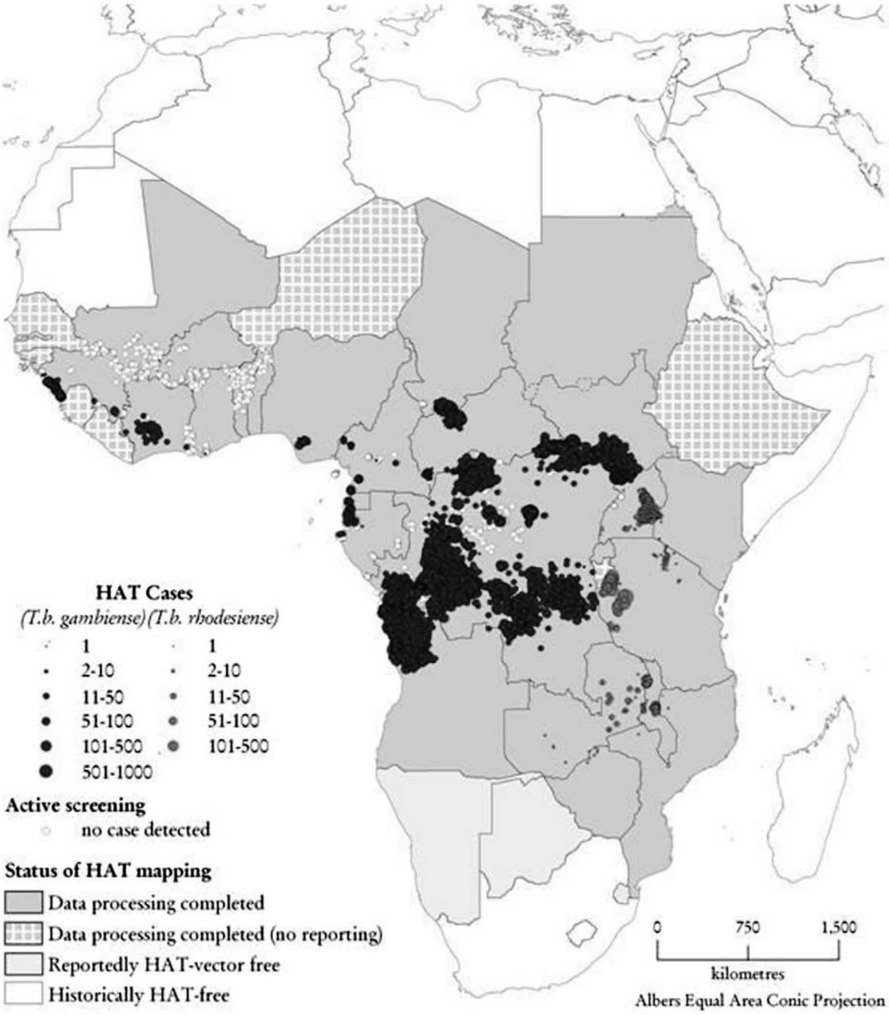

HAT Cases
(T.b. gambiense) (T.b. rhodesiense)
1
2-10
11-50
51-100
101-500
501-1000
1
2-10
11-50
51-100
101-500
Active screening
no case detected
Status of HAT mapping
Data processing completed
Data processing completed (no reporting)
Reportedly HAT-vector free
Historically HAT-free
0
750
1,500
kilometres
Albers Equal Area Conic Projection

FIGURE 39.1 **Distribution of HAT cases in Africa.** *Reproduced, with permission from WHO (2013). Control and surveillance of human African trypanosomiasis. TRS 984;* [*in press*].

T. b. gambiense, representing 97% of the reported cases [7]. Although currently rare, *T. b. rhodesiense* tends to emerge in large epidemics (e.g., more than 250,000 people were estimated to have been killed in Uganda between 1900 and 1915) [8].

T. b. gambiense and *T. b. rhodesiense* are transmitted through the bite of an infected tsetse fly (Figure 39.2). Congenital transmission and transmission through direct blood contact have been reported, but rarely [1,2]. Tsetse flies of the genus *Glossina* are commonly divided into the fusca group, which inhabits forested locations, the morsitans group, which inhabits savannah, and the palpalis group, which inhabits riverine/forest habitats [9]. *T. b. gambiense* are most commonly transmitted by palpalis group flies (*G. fuscipes* and *G. palpalis*). Unlike in malaria, where only female mosquitoes transmit *Plasmodium* parasites, both male and female tsetse flies feed from humans and can transmit trypanosomes. It takes 18–35 days for trypanosomes to complete their life cycle within the tsetse fly, emerging in the fly's salivary glands as human-infective metacyclic forms, after having been through several life-cycle forms within the insect.

In addition to infection of humans, the parasites can infect other mammals. In the case of *T. b rhodesiense*, the animal reservoir is an important part of the epidemiology and the disease is zoonotic. *T. b. gambiense* was classically believed not to infect animals other than man; however, sensitive molecular tools have, in recent years, shown these parasites to be capable of infecting animals, although the role of these reservoirs in disease epidemiology is not known [10].

The prevalence of both gambiense and rhodesiense HAT has been declining in the twenty-first century, and fewer than 7000 cases were reported in 2011 [11]. However, the epidemic nature of the disease was associated with several periods of resurgence and remission over the twentieth century. For example, gambiense disease was estimated to have killed nearly a million people in equatorial Africa over the years 1896–1906 [12]. It has yet to be established which factors influence the ebb and flow of HAT epidemiology,

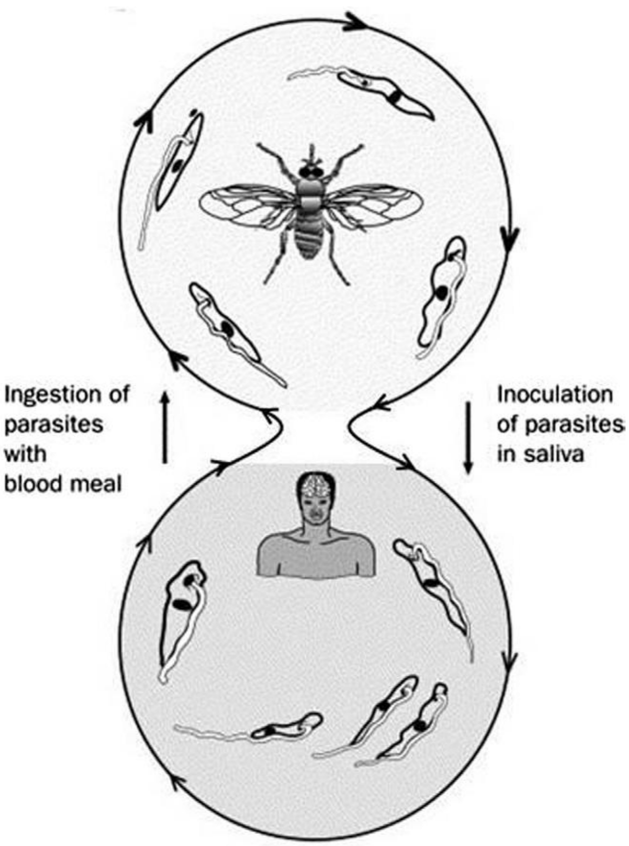

Ingestion of parasites with blood meal
Inoculation of parasites in saliva

FIGURE 39.2 **Life cycle of *Trypanosoma brucei* subspecies.** *T. brucei* parasites proliferate in the mammalian bloodstream as long slender forms, which can also enter the central nervous system. Some parasites differentiate into a nonreplicative short, stumpy form that has altered its biochemical physiology to enable survival in the midgut of the tsetse fly when taken up in a blood meal. Within the tsetse midgut, the parasites differentiate into a replicative procyclic form and in this form migrate to the fly's salivary gland, where they ultimately differentiate, via several other stages on route, into the metacyclic form that is preadapted to survive when taken back into the human host. *Reproduced, with permission from WHO (2013). Control and surveillance of human African trypanosomiasis. TRS 984;* [*in press*].

although an inverse relationship between surveillance efforts and disease prevalence is apparent. There is currently optimism that our improved understanding of the disease's epidemiology, the availability of drugs and pathways for their distribution, and other intervention tools, such as tsetse fly control, might enable HAT to be eliminated as a public health problem by 2020 and eliminated (fewer than one case per 10,000 people in endemic regions) by 2030 [13].

39.2 THE GENETICS OF TRYPANOSOME LYTIC FACTORS AND PARASITE FACTORS ENABLING PROTECTION AGAINST LYSIS

The African trypanosomes branched early in eukaryotic evolution and are classified within the phylogenetic order Kinetoplastida in the Protozoan subkingdom [2]. African trypanosomes belong to the salivarian section of the genus *Trypanosoma*, being injected into mammalian blood via the bite of a tsetse fly (or other biting flies for some nonhuman infective trypanosome species). The stercorarian section describes parasites that include the American trypanosome, *Trypanosoma cruzi*, that is transmitted in the feces of its reduviid bug vector. *T. cruzi* causes Chagas disease, endemic across large parts of Latin America [14], but it is not discussed in this chapter, which confines itself to human African trypanosomiasis.

T. b. gambiense and *T. b. rhodesiense* are identical by microscopical analysis (Figure 39.3). However, in recent years, specific molecular markers have emerged for these two subspecies. The so-called serum resistance-associated (*SRA*) gene [15,16] is ubiquitous in *T. b. rhodesiense* isolates, while type 1 *T. b. gambiense* isolates possess a distinctive gene called *TGSGP* [17]. A second *T. b. gambiense* lineage, defined as human infectious trypanosomes found in West and Central Africa but lacking the *TGSGP* gene, was defined as type 2 *T. b. gambiense*, but no representatives of this group have been isolated in over 20 years; for this reason, their relevance today is uncertain [18]. Both *SRA* and *TGSGP* play roles in protecting trypanosomes against trypanosome lytic factors found in human serum.

The best understood of these lytic factors is apolipoprotein L1 (APOL1), which is part of the high-density lipoprotein fraction of plasma, long known to be a primary determinant of trypanosome lysis [19]. Particles carrying APOL1, along with a haptoglobin-related protein that is recognized as a ligand by a specific haptoglobin-binding receptor (HPR), brings this complex into trypanosomes by receptor-mediated endocytosis [20,21]. It appears that when APOL1 is released from the complex it can bind to and permeabilize the parasite's endosomal membrane, causing ion imbalances, cell swelling, and death of the parasites. The *SRA* gene product in *T. b. rhodesiense* binds to, and thus neutralizes, APOL1, rendering these parasites resistant to human serum [22].

SRA appears to be a truncated derivative of a variant surface glycoprotein (*VSG*) gene, probably generated by a rare mutation but subsequently selected in parasite lines that gained human infectivity as a trait [23]. The mechanism by which TGSGP inhibits the activity of APOL1 has yet to be characterized. It has also been proposed that *T. b. gambiense* internalizes trypanosome lytic factors less readily than other species because of lowered expression of the HPR [24] and because changes to membrane fluidity might affect uptake. A second trypanolytic factor is less well understood but also critical in nonimmune-mediated trypanolysis [20].

The central role of APOL1 in trypanolysis was confirmed when an individual in India was shown to be infected with *Trypanosoma evansi*, a parasite related to *T. brucei* subspecies but with a key biological distinction [5,25]. *T. evansi* do not depend on the tsetse fly for transmission, instead being transmitted mechanically between mammalian hosts by other biting flies such as tabanids,

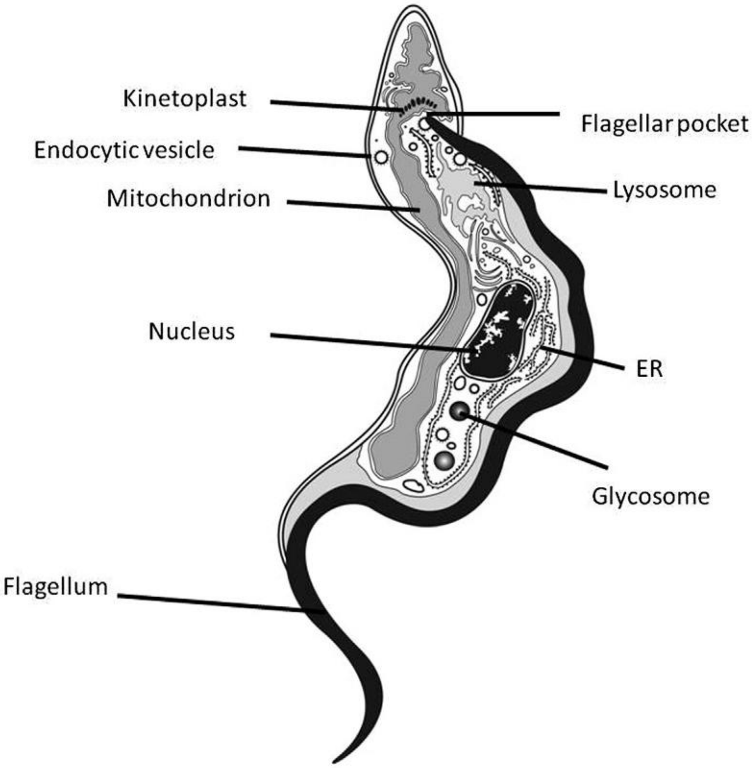


FIGURE 39.3 **Morphology of *Trypanosoma brucei*.** Trypanosomes are protozoa, around 20 micrometers in length and 5 micrometers in width. They are characterized by several organelles, including a flagellum that protrudes from the front of the cell. The single mitochondrion contains an intercatenated network of circular DNA molecules called the kinetoplast. Endocytosis can only occur within the region of the cell from which the flagellum emerges, called the flagellar pocket. This is the only part of the cell not defined by a subpellicular microtubule array. Other unusual organelles include glycosomes, peroxisome-related organelles in which glycolysis occurs, and acidocalcisomes, which serve as repositories for intracellular cations and might play roles in cellular pH and volume homeostasis.

including horse flys [26]. Tsetse-transmitted trypanosomes undertake a proliferative life cycle within the fly, and their biochemical physiology is adapted to enable them to use substrates (e.g., proline) as key energy and carbon sources [27]. For this purpose, the mitochondrial metabolism that is repressed during life in the mammalian bloodstream is derepressed in the insect stage. Here in the tsetse fly, a partial TCA cycle and electron transport chain operate to provide sufficient energy to the parasites to complete this life cycle [28]. Having lost a dependency on tsetse transmission, *T. evansi* has also lost a dependency on these key aspects of mitochondrial metabolism and in many instances has also lost its mitochondrial genome, known as the kinetoplast [29]. The loss of dependency on tsetse flies for transmission has enabled *T. evansi* to extend its geographical range beyond Sub-Saharan Africa, reaching Latin America and Asia. *T. evansi* does not, however, have any molecular mechanisms allowing avoidance of APOL1 or other trypanosome lytic factors. Therefore, it is restricted to nonprimate mammalian species, representing an important veterinary health risk in many parts of the world.

In 2005 an Indian farmer presented with fever and malaise and blood tests revealed a fulminating *T. evansi* infection that was successfully treated with suramin [5]. The same patient was later shown to be defective in APOL1 [25]. Other cases have come to light since then, but the true extent of infection of APOL1-defective individuals with *T. evansi*, or other trypanosome species that are not normally infectious to man, is not known.

The APOL1 story has recently taken another intriguing twist. As genetic data accumulated in the United States on the prevalence of kidney disease, including end-stage kidney disease (ESKD) and HIV-associated nephropathy (HIVAN), the presence of heritable susceptibility in African American and, to a lesser degree, Hispanic populations in the United States became clear [30]. Genome-wide association studies (GWASs) using admixture mapping to seek ancestral associations showed that a region on chromosome 22q12.3 is associated with chronic kidney disease [31]. The associated locus contains 30 genes, including the *APOL1* gene. The over-representation in African Americans is proposed to derive from the forced emigration of people from West Africa as part of the historical slave trade. Common variants, designated G1 and G2, found in these populations have been proposed to allow lysis of SRA-expressing *T. b. rhodesiense* that resist lysis by other APOL1 variants. In the case of the G2 variant, this relates to a lessened ability to bind to SRA, although the reason for the G1 variant not lysing is uncertain given that it continues to bind to SRA. Neither variant lyses *T. b. gambiense*, the trypanosome species endemic to West and Central Africa today [32].

Based on these observations, it has been proposed that African trypanosomes expressing a protein similar to SRA might have been a strong selective force in West Africa at some point in the evolution of human populations. This would have led to the evolution of the G1 and G2 alleles of *APOL1* that remain prevalent there today. *T. b. gambiense* then would appear to have emerged later, having the ability to infect individuals with the G1 and G2 alleles (its ability to resist lysis being independent of SRA). This theory is at odds with the notion that *T. b. rhodesiense* arose relatively recently (perhaps as recently as the early twentieth century in East Africa) [8], although a plausible scenario is that ancestral trypanosome populations in West Africa with an *SRA*-type gene were later lost as human populations were selected with the resistance alleles. The current *T. b. rhodesiense* subspecies could have emerged independently in East Africa at a later date.

The discovery of allelic variants of genes that appear to have emerged in human populations because of parasite-imposed selection pressures is not restricted to trypanosomes. *Plasmodium* parasites that cause malaria have been known for many years to have selected various traits through gene variants whose protection against malaria outweighs the deleterious effects associated with those variants [33]. Inheritance of sickle cell anemia, through the protective effects of sickle cell trait in heterozygous individuals, is one example [34]. The other best known example has been the selection of variants of glucose-6-phosphate dehydrogenase (G6PD) where reduced levels of that enzyme lead to reduced production of cellular NADPH, which in turn is critical to maintenance of cellular redox balance through glutathione [35]. The widespread prevalence of G6PD deficiency was noted in African Americans treated with primaquine, an oxidative stress–inducing antimalarial agent. Ironically, the diminished ability to deal with oxidative stress provides protection of red blood cells against infection with malaria parasites, and various *G6PD* alleles were selected in evolution in malarious regions, including the *G6PDA−* allele in Africa and the *G6PD* Mediterranean allele in the Middle East. Only later, when oxidative stress–inducing drugs like primaquine were used, did the linkage come to light [36].

To date, the geographical prevalence of *APOL1* alleles that protect against *SRA*-bearing trypanosomes and the association of those alleles with kidney disease have not been investigated for other associations with enhanced toxicity when treating with trypanocidal drugs. However, it is noteworthy that both suramin and pentamidine have nephrotoxic liabilities; thus, it would be of interest to determine whether the distribution of host responses to those drugs in any way related to *APOL1* alleles.

The APOL1 story has emerged as a key paradigm in the so-called evolutionary arms race between host and parasite. At the level of the parasite, apart from genes such as *SRA* and *TGSGP*, which have clear effects on infectivity, other genes affect virulence. For example, although the classical description of *T. b. rhodesiense* infection indicates a rapidly progressing, virulent disease, in Malawi this subspecies is associated with a more insidious progression, similar to that of classical *T. b. gambiense* [37,38]. Conversely, there are reports of rapidly progressing *T. b. gambiense* [39] along with the textbook chronic form. There are also cases of gambiense infections where patients have never reported any symptoms of HAT and yet carry the parasites [40,41]; these individuals may represent a cohort of asymptomatic carriers. Sensitive PCR-based diagnostic tools are making it easiser to identify such cases. Their role in the epidemiology of the disease may be of great importance, considering the WHO goal of HAT elimination, as they may act as long-term reservoirs. Ethical issues pertaining to the treatment of such asymptomatic cases are lessened with the availability of nifurtimox-eflornithine combination therapy (NECT), which is much safer than the highly toxic melarsoprol that would have previously been required to treat stage 2 cases. Several reports of gambiense parasites adapting for growth in rodents have shown a range of virulence types [42,43], although parasite genes controlling this variation in virulence are not known.

39.3 IMMUNE RESPONSES AND IMMUNOAVOIDANCE

In addition to the innate immune system, adaptive immunity also plays a role in HAT progression. Immunological responses to trypanosomes have been studied in rodent models, where it has been shown that B and T cells along with macrophages play key roles in antitrypanosome responses [44]. Type 1 cytokines (IFN-γ and TNF-α) accompany production of nitric oxide (NO), which has a trypanotoxic effect. Interestingly, trypanosomes appear to have the ability to stimulate host arginase, which converts arginine to ornithine and thus diverts arginine from NO production [45]. In humans, cytokine analysis has suggested that IL6 and IL8 are produced in stage 1 infections while IL10 is significantly elevated in stage 2 patients [46], although inter-study variability has made it difficult to provide definitive trends [39].

Nevertheless, other works have suggested using cytokine profiles as diagnostic biomarkers to assess the disease (see later). Rodent models (and cattle models) have allowed genetic studies to seek quantitative trait loci (QTLs) that have an impact on parasite progression in African trypanosomiasis [47,48]. In mice infected with the veterinary trypanosome *T. congolense*, an association was shown between infection and loci on chromosomes 17, 5, and 1. These loci were designated *Tir*1, *Tir*2, and *Tir*3. Although individual genes that are associated with infectivity have not been proven, within these loci are genes encoding TNFα (*Tir*1), IL10 and a regulator of its expression called *CYPr2* (in *Tir*3b), and another IL10 regulatory factor called *CYPr3* in *Tir*2. In humans, analysis of polymorphisms in cytokine genes has shown a weak association to particular IL6 and IL10 variants while IL1 and TNFα have been proposed as useful contributors to possible diagnostic tools [49].

The demonstration of associations between molecules of the adaptive immune response and trypanosome infection shows how parasites must avoid the adaptive immune system. Trypanosomes are immunogenic and stimulate antibodies of the IgM and IgG classes [50]. A dense coat (~15 nm in thickness), comprising dimerized variant surface glycoproteins (VSGs), excludes antibodies and complement proteins from the plasma membrane, where various vital functional proteins, including the transporters involved in nutrient acquisition and various communication-related processes, are located. Antibodies aimed at the immunogenic coat proteins kill parasites. However, African trypanosomes possess up to 1000 individual *VSG* genes. They express these one at a time, more commonly by duplicating basic copies of the *VSG* genes and transposing them to telomeric expression sites of which only one can be activated at a time [51–53].

Many of the genes are actually pseudogenes in their nascent state, but recombination events can recombine parts of multiple genes to create new variants. Parasites expressing one VSG type are killed by the antibodies raised to that particular type, but a small number of parasites then express a separate VSG and change to a new coat. These variants then proliferate until antibodies to this VSG emerge, but this serves only to select another VSG switch variant. In this way the parasite population can survive immunological assault more or less *ad infinitum*.

Trypanosomiasis is also associated with general immunosuppression that also enhances the parasite's ability to survive [54,55]. The so-called trypanosome-suppressive immunomodulatory factor (TSIF) [56] has been identified as a mediator of this immunosuppression, although more details on how it achieves this effect are required.

39.4 PATHOLOGIES ASSOCIATED WITH HAT

The pathology of human African trypanosomiasis depends on the causative species and the stage of the disease. While the general symptoms are similar for both HAT types, the onset and severity are divergent. Rhodesiense HAT progresses to stage 2 within weeks and death normally happens within 6 months of infection. Gambiense HAT, on the other hand, takes around 18 months to reach second stage and a further 18 months, or more, prior to death [57]. Interestingly, a number of instances have been reported where disease progression of gambiense infection can be much slower; indeed, as mentioned above, asymptomatic carriers have been described [40]. However, little information exists that relates the host or parasite factors that lead to these variations in disease progression.

After infection, through the tsetse bite, the disease enters its first hemolymphatic stage, with parasites confined to the blood and lymph systems [58]. The second, or meningoencephalitic, stage involves invasion of the central nervous system (CNS) by the trypanosomes. While stage 1 is generally accompanied by nonspecific symptoms such as headache and malaise, stage 2 is characterized by progressive neurological breakdown, including the change in sleep/wake patterns that led to the disease being referred to as sleeping sickness [59].

For gambiense HAT, a trypanosome chancre at the site of the tsetse fly bite is only rarely observed, while it is common with rhodesiense HAT. The first stage is then accompanied by fever, headache, anemia, pruritus, lymphadenopathy, weakness, and asthenia. These symptoms seldom prompt a person to seek health services. A lymphadenopathy in the cervical lymph nodes is typical for gambiense disease. It was classically termed "Winterbottom's sign" after the British physician Thomas Winterbottom and used by slave traders to distinguish those likely to die during shipment to the Americas. This lymphadenopathy is seldom seen in rhodesiense infections. Pruritus, too, is seen in only a cohort of patients, but to date the reasons that populations show variability in this complaint are not known, although coinfection with different helminthiases (e.g., loiasis or onchocerciasis) might provoke different responses in different individuals [60]. Cardiac involvement is common, with ECG alterations recorded in around 50% of patients in stage 1 and 70% in stage 2 gambiense HAT [61], although clinical heart failure is rare.

Somnography has been used to reveal dysregulation of circadian rhythms in the sleep/wake cycle and fragmented sleep, with sleep-onset rapid-eye-movement (SOREM) episodes being characteristic of the HAT-associated sleep disorder [62,63]. Various other nonspecific neurological changes occur in stage 2 HAT, including headaches, mood and behavioral changes, irritability, aggressive behavior, hyperactive or apathetic attitudes, indifference, depression, and/or delirium with hallucinations. Peripheral neuropathy is also evident, and tone disorder with abnormal movements is common. Motor weakness, tremor, bradykinesis, and a Parkinson-like rigidity are also observed [58,59]. Most of these disorders are reversed on treatment with trypanocidal drugs.

In very advanced cases, dementia and epilepsy are frequent, and incontinence, malnutrition, and cachexia can become evident. Patients can fall into a coma prior to death.

Dysfunction of the endocrine system is frequent with thyroid and adrenocortical function compromised [64]. Estrodiol levels are decreased in females and testosterone levels in males. Sexual impotence is often seen [65].

HIV coinfection influences the outcome of many parasitic and other microbial diseases. However, there is no evidence that HIV prevalence is higher in HAT patients than in controls [66], although there is some indication that immunosuppression might influence treatment outcomes. For example, in one study, 14 HIV-negative patients treated with melarsoprol left the hospital cured, while 4 HIV-positive patients in the same study suffered relapse [67], although further evidence is lacking.

Interestingly, when visitors to endemic regions become infected with trypanosomes, they appear to develop symptoms quickly, irrespective of the infecting subspecies. Fever appears rapidly, and the trypanosomal chancre is seen in most rhodesiense patients and in around half of those infected with gambiense [68]. Most travelers seek medical help prior to stage 2 onset; hence, little is known about this stage in individuals from nonendemic countries. The causes of different disease progression in nonindigenous visitors to HAT endemic regions compared to inhabitants of those areas are not known, but genetics or nutritional status might play a role.

39.5 DIAGNOSIS OF HAT

Because HAT is a lethal disease, diagnosis is vital. Different drugs are required to treat the disease depending on the causative subspecies and also on the stage of disease. Therefore, it is necessary to diagnose the presence of parasites, to identify the causative subspecies, and to stage the disease according to whether it is in the first, hemolymphatic, or in the second, neuroencephalitic, stage.

In stage 1 the general nature of symptoms makes symptomatology-based diagnosis difficult. Furthermore, given the dangers associated with current drugs (see later), treatment cannot begin without positive diagnosis, which generally requires identification of parasite by microscopy [4].

A confirmed case is defined as one in which parasites have been identified in blood or CSF. A suspected case is one in which antibodies to parasites are noted or parasite nucleic acids are detected, but parasites have not been found by microscopy. Currently, since *T. b. gambiense* and *T. b. rhodesiense* have distinct geographical localizations, identifying trypanosomes in people from a region where one or the other parasite is known to be present can be enough to speciate (unless patients have traveled to regions endemic for the other parasites), although molecular markers are now available to provide precise characterization.

For *T. b. gambiense* infections, an antibody test, usually in the format of card agglutination, is available [69]. This assay was made possible because most gambiense parasites express particular variant surface glycoproteins during early stages of infection; these stimulate antibody production, which can be detected through their binding to the cognate antigens. The same VSGs are not generally present in *T. b. rhodesiense*; hence, these serological tests cannot be applied to rhodesiense infections, and presently diagnosis of this subspecies starts immediately with microscopic investigation of blood.

In the card agglutination test for trypanosomiasis (CATT), trypanosomes expressing the LiTat1.3 antigen are cultivated, freeze-dried, and then used as an agglutination source on cards, where they are mixed with blood from a suspected patient's finger prick [69]. The freeze-dried material agglutinates if antibodies specific to the parasite are present. Whole blood, undiluted, is most easily used, but plasma in serial dilutions (up to 32-fold diluted or more) can offer better specificity [4]. The CATT can be performed on blood-impregnated filter paper for delayed testing, with the drawback of reducing sensitivity.

Recently, the same antigens were purified and incorporated into dipstick-based tests [70], where patient serum can interact with the antigen on a solid state matrix and create a color change to a strip in the event of antigen-antibody forming. The fact that 3–4 weeks must elapse prior to seroconversion to the salient antigens hampers the approach in early diagnosis. Moreover, there are *T. b. gambiense* strains that lack the LiTat1.3 antigen [71] and therefore do not cause seroconversion to this test antigen. It is also possible to determine whether patient serum can induce lysis of trypanosomes, which will indicate the presence of those trypanosome-specific antibodies. Indirect immunofluorescence can be used, by applying trypanosomes to glass slides and then adding patient serum and secondary anti-human IgG antibodies with fluorescent labels [72]. ELISA tests based on the same basic antigen–antibody interaction are also available [73]. However, such tests are more expensive and require fluorescence microscopy; thus, they are not widely used.

The serological tests have received widespread use in mass screening campaigns, and this has underpinned some of the advances in bringing the disease under control in the twenty-first century. However, their fallibility means that parasite identification by microscopy has remained the diagnostic gold standard. Since trypanosomes are highly motile, inspection of wet blood smears can be the best means of detection. Lymph node aspirate or CSF from a lumbar puncture can be used as well.

The very low parasitemia associated with *T. b. gambiense* infections offers a particular challenge to microscopic evaluation. With fewer than 1000 parasites per mL of blood in many instances, microscopy can require great patience and skill to find just a single trypanosome. Methods for concentrating parasites in blood are therefore used where possible.

When enlarged cervical lymph nodes are present in early stages of infection, aspirate microscopy can be easier than seeking parasites in blood [4]. However, because of the nonuniformity in presentation with enlarged lymph nodes, this cannot be routinely used. For this reason, blood films are prepared. These can be either thin films, where parasite motility can be used to assist identification, or Giemsa-stained thick films, where the trypanosome's unique shape and staining of the nucleus and kinetoplast characterize the parasites. The development of relatively cheap and robust fluorescence microscopes that use light-emitting diodes (LEDs) [74] and the fact that various fluorescent stains, including acridine orange, accumulate in trypanosomes [75] have recently been shown to facilitate microscopical identification of trypanosomes in blood.

Enrichment of parasites from blood samples also improves sensitivity. For example, lysis of red blood cells by ammonium chloride followed by centrifugation and examination by microscopy of a thin or thick smear of the sediment (stained with Giemsa or acridine orange) improves sensitivity [75]. Other parasite concentration tests include the Woo test, a microhematocrit centrifugation technique (mHCT) [76], also called the capillary tube centrifugation technique (CTC), which uses capillary tubes sealed with plasticine that are three-quarters filled with blood from a finger prick. These are centrifuged at 12,000 × g for 5 minutes, after which parasites, along with white blood cells, concentrate at the layer between the erythrocytes and plasma, allowing targeted microscopy of this region. The technique allows identification of parasites in infections as low as 500 trypanosomes/mL.

A related approach, the quantitative buffy coat technique, involves adding parasites to a capillary tube that has been precoated with EDTA and acridine orange. After centrifugation, parasites can be identified by virtue of having taken up the acridine orange, which can be detected readily with UV fluorescence [77]. This technique is a little more sensitive than the Woo test, but it requires a fluorescence light source and a supply of the acridine orange–coated capillary tubes.

The most widely used parasite concentration technique is mini-anion exchange centrifugation [78]. This technique works because the trypanosome's surface coat is neutral in buffered solutions at pH 8 while the phospholipids of red cells leave them negatively charged. The anion exchange matrix diethylaminoethyl (DEAE) cellulose then binds red cells to the matrix, while the parasites elute from the column. Columns are manufactured to enable application of 100 or more microliters of blood, and the eluted parasites are captured in a pipette that can be centrifuged and viewed by microscopy. The use of precentrifuged blood, with volumes up to 5 mL, allows to capture parasites in a buffy coat between the erythrocyte and plasma layers, improving the sensitivity to as few as 10 parasites per mL of blood.

Molecular methods that detect parasite DNA are of increasing importance to diagnostic algorithms. The ability to detect frozen, or filter paper–dried, samples increases their attractiveness. Multicopy DNA sequences specific to the *T. brucei* group are very sensitive [79] and can be combined with the highly specific (but lower copy number and thus less sensitive) subspecies-defining probes: *SRA* for *T. b. rhodesiense* and *TGSGP* for *T. b. gambiense* [80].

The requirement for a thermocycler and electricity supply can hinder the use of PCR in regions where HAT is endemic. For this reason, the loop-mediated isothermal amplification (LAMP) approach has been developed [81], which allows amplification of target DNA without a need for thermal cycling. *Trypanozoon*-specific multicopy DNA sequences and the *T. b. gambiense*-specific *TGSGP* gene [81] have both been used in LAMP assays, which use the dye hydroxynapthol blue in detection of double-stranded DNA produced during the reaction.

Microscopical approaches can be easier for rhodesiense infections than for gambiense infections, since parasitemias are generally higher in the former. The relatively frequent development of a chancre at the site of the tsetse bite means that this can be biopsied and examined microscopically for parasites soon after infection. Otherwise, thin and thick film blood smears can be prepared, using parasite motility to aid identification in thin smears and Giemsa staining to identify parasites in thick smears. DNA-based tests, including PCR and the LAMP method, are equally applicable to rhodesiense, with the *SRA* gene serving as a subspecies-dependent marker.

39.6 STAGING THE DISEASE

In stratifying patients for treatment, determining the stage to which the disease has progressed is vital. The drugs used to treat stage 1 and stage 2 disease differ and, since stage 2 drugs generally are more difficult to administer and dangerous, they should only be given to the relevant stage.

Currently, staging depends upon analysis of CSF obtained by lumbar puncture. Identification of trypanosomes in CSF is considered diagnostic of stage 2 disease, although in some cases low parasite numbers mean that high white cell count (higher than 5 cells per microliter or 20 in some endemic countries) can be taken as diagnostic [4,82]. For many years cases where no parasites were found in CSF, but 5–20 cells were found, were considered to be "early second stage," and recommendations to treat with pentamidine, which appeared curative [83], were made, given the exceedingly toxic nature of melarsoprol. Now, however, nifurtimox-eflornithine combination therapy (see below) is preferred for stage 2 disease; since this therapeutic choice is relatively safe, it is also given to patients in the early second stage. Generally, even in the absence of parasites, patients with >20 WBC per microliter are judged to be in stage 2 and treated accordingly.

Because of the difficulties in finding parasites in CSF and the imprecise utility of white cell counts, efforts have been made to look for other surrogate markers. Protein concentration was considered to have potential [84]; however, the low sensitivity precluded widespread application. In recent years, more specific tests were proposed for IgM [85], and other protein biomarkers were considered.

Biomarker research has sought both proteomic and metabolomic markers that might diagnose HAT and also offer stage prediction without needing to find parasites. In one study, quantification of a range of known inflammation-related proteins were determined in CSF of *T. b. gambiense* patients [86]. CXCL10 most accurately distinguished stage 1 and stage 2 patients with a sensitivity of 84% and a specificity of 100%. Moreover, using a panel of proteins including CXCL10, CXCL8, and H-FABP improved the detection of stage 2 patients to 97% sensitivity. Further analysis using 2-dimensional gel electrophoresis to identify quantitative differences in protein levels between stage 1 and stage 2 patients showed that levels of osteopontin and β-2-microglobulin could also stage the disease with sensitivities of 68% and 78%, respectively, and 100% specificity [87]. Combining the two markers increased sensitivity to 91%. A third study investigated another panel of proteins and showed that ICAM-1 and MMP-9 accurately discriminated between stage 1 and stage 2 HAT patients (95% sensitivity, 100% specificity) [88]. The inflammatory metabolite neopterin was also measured and was shown to predict stage 2 disease [89]. To date, a combination of all of these antigens has not been tested, although it should be interesting to see how sensitivity and specificity perform in a panel of the full set. No successful efforts have yet been published reporting markers derived from untargeted approaches in blood or urine that might predict the stage of the disease without requiring invasive lumbar puncture. As methods in proteomics and metabolomics become more sensitive, however, it seems likely that better functional biomarkers will emerge.

The identification of the metabolite neopterin indicates that metabolomics offers a potential route to biomarker discovery [89]. Early studies showed that *T. b. gambiense* metabolizes tryptophan, tyrosine, and phenylalanine to indole 3-pyruvic acid, 4-hydroxphenylpyruvic acid, and phenylpyruvic acid, respectively, *in vitro*, and that rodents infected with *T. b. gambiense* also secreted catabolites of these aromatic amino acids [90,91]. ^{1}H NMR was also used to probe metabolic responses in *T. brucei* infection in mouse urine and plasma over seven days. Plasma and urine of infected animals showed elevated lactate levels, and plasma from infected mice also showed depletion of glutamine, leucine, isoleucine, and valine [92]. In urine, there were increases in the branched chain amino acid metabolites 3-methyl-2-oxovalerate and 2-oxoisovalerate, as well as D-3-hydroxybutyrate. Lipid concentrations varied, and diminished quantities of hippurate were reported along with increases in trimethylamine and 4-hydroxyphenylacetic acid. The major distinguishing metabolites were the elevated levels of urinary 3-methyl-2-oxovalerate, 3-carboxy-2-methyl-3-oxopropanamine, and 4-hydroxyphenyl acetic acid [92]. A potential drawback in translating information relating to heavy infections in rodents to very low parasitemia situations in humans relates to whether these biomarkers are proportional to parasitemia; if so, levels may fall below the detection limit in humans. LC–MS offers enhanced sensitivity for quantification of specified metabolites [93], and work is under way to determine whether markers for infection and stage 2 involvement can be detected using these platforms.

As described in the next section, HAT treatment is stratified based on the causative parasite and the stage of the disease. No treatments are 100% efficacious, and reports of resistance to current drugs have increased in recent years. Therefore, it is necessary to assess treatment outcome through post-treatment follow-up. Recent WHO recommendations suggest that follow-up of patients treated for stage 1 is necessary only if patients feel symptoms indicative of relapse (other than in clinical trials, when mandatory follow-up, six monthly for 18–24 months, is required).

For stage 2 HAT, however, the difficulties associated with cure are such that patients should be followed up every six months for two years. If parasites cannot be found and CSF white cell counts are below 5 cells per microliter, patients can be considered cured. If white cell counts are greater than 50 per microliter, these patients should be considered a failure and retreated. If white cell counts are 6–50 cells per microliter, patients should be re-evaluated at 12 months. Guidelines for rhodesiense follow-up are not available but, given the more rapid progression, a similar regimen, but with shorter time frames, should be considered.

39.7 TREATMENT

39.7.1 Current Drugs

The history of chemotherapy is entwined with the history of HAT therapy [94]. The pioneering German physician/scientist Paul Ehrlich experimented with various agents to kill trypanosomes. His research led directly to the introduction of napthalene dyes with trypanocidal activity, themselves the precursors of the suramin still used today, and also the organic arsenicals, which served as precursors of melarsoprol, which remains the only drug available to treat stage 2 rhodesiense disease.

Today five compounds are used in the treatment of HAT, depending on the causative subspecies and the stage

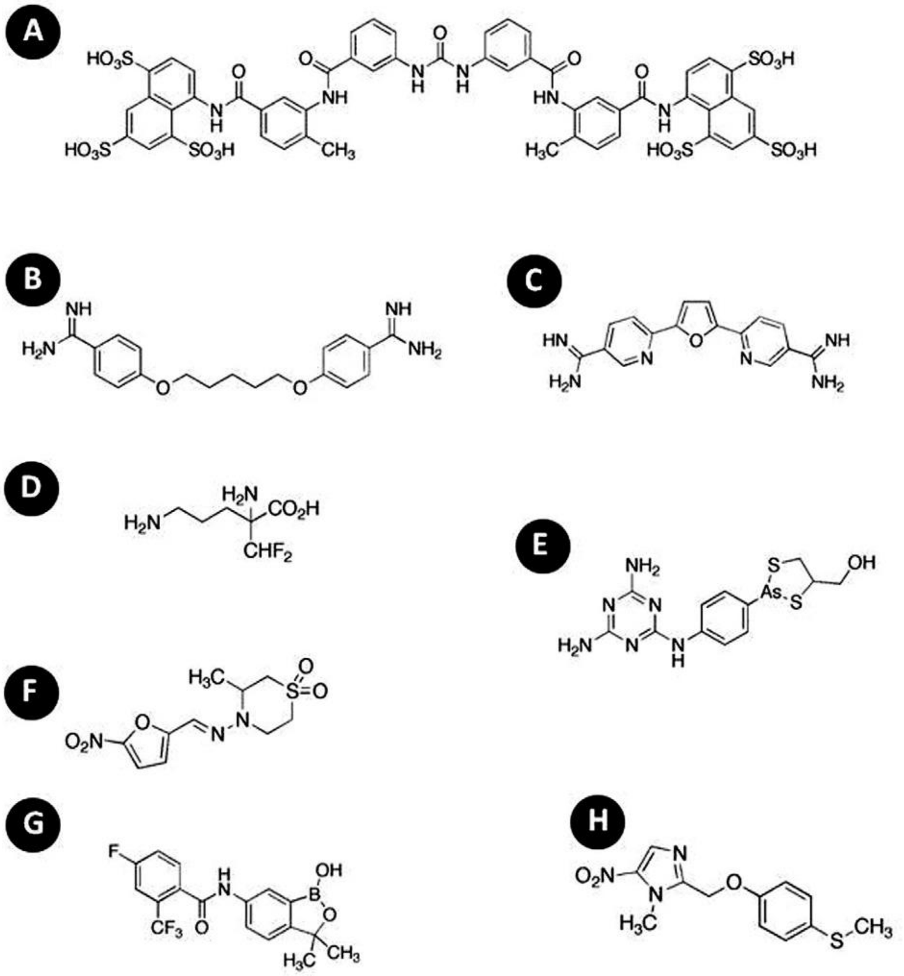


FIGURE 39.4 **Structures of current drugs and compounds in HAT trials.** (a) suramin; (b) pentamidine, (c) CPD-0801; (d) eflornithine; (e) melarsoprol; (f) fexinidazole; (g) SCYX-7158; (h) nifurtimox.

of the disease. For gambiense HAT in stage 1, pentamidine is used, while in stage 2 the preferred treatment today is nifurtimox-eflornithine combination therapy (NECT), although eflornithine alone, or occasionally melarsoprol, are still used. For rhodesiense disease in stage 1, suramin is the drug of choice, while in stage 2 melarsoprol is used. The twenty-first century has witnessed the emergence of new paradigms in the discovery and development of new drugs for neglected and other infectious diseases, including HAT.

Most of the available HAT treatments were introduced many years ago, without the kinds of evaluation required for drugs being trialed in the modern era. Suramin, for example, was introduced in the 1920s, pentamidine in the 1930s, melarsoprol in the 1940s, nifurtimox (for Chagas disease) in the 1970s, and eflornithine in the 1980s. Given that all available drugs suffer various drawbacks—the need for protracted parenteral administration, insecure manufacture and distribution, serious adverse events in some cases, and increasing incidence of treatment failure—a robust pipeline for new medications is essential. Structures of current drugs and compounds in late preclinical or clinical trials are given in Figure 39.4.

39.7.2 Stage 1

39.7.2.1 *Pentamidine*

Pentamidine is used to treat stage 1 *T. b. gambiense* disease [95]. Administration is by intramuscular injection once daily at 4 mg kg^{-1} over seven days (intravenous injection provokes a transient hypoglycemia). The drug is highly trypanocidal; IC_{50} values are in the order of 1–10 nM in a typical 3-day drug sensitivity assay in rich medium, falling to even lower levels in a novel minimal medium developed using a metabolomics approach that showed how most components of the classical rich medium were unnecessary for parasite growth [96]. Pentamidine is concentrated to very high levels in trypanosomes using several transporter systems, including the P2 amino-purine transporter and a high-affinity pentamidine transporter (HAPT1) that has recently been shown to be encoded by the aquaglyceroporin-type transporter (*AQP2*) [97]. A further low-affinity pentamidine transporter (LAPT1) is also involved in uptake [98]. How pentamidine exerts its activity within cells is unknown, although it is possible that activity is multifactorial through concentration to high levels. Metabolomics experiments failed to identify a clear single enzymatic system affected by the drug [99]. Diamidines bind avidly to DNA, and mitochondrial dysfunction has been associated with diamidine treatment.

Extensive tissue retention and binding to serum proteins contribute to a large volume of distribution and a long terminal half-life. Since the drug is positively charged at physiological pH, it cannot diffuse across membranes, precluding oral availability as well as brain permeation. Pentamidine is also metabolized to a substantial degree by cytochrome P450 systems [95]. Although it has been used for many years, pentamidine can be toxic. Injection-related pain, nephrotoxicity, leucopenia, and liver enzyme abnormalities are all common, as is the risk of hypoglycemia. Of note is that the clinical trials of DB289, a methoxy prodrug precursor to the diamidine DB75, was abandoned when serious incidence of nephrotoxicity accompanied anticipated hepatotoxic effects [100]. Further development of other diamidines is now dependent on an early and clear understanding of likely nephrotoxicity and hepatotoxicity. Several aza analogs of DB75 have shown the remarkable capability of clearing stage 2 infections in rodent models, which indicates uptake into the brain, presumably via specific transporters; these compounds have been shown to be substrates of organic cation transporters in brain endothelia [101].

39.7.2.2 Suramin

Suramin is used in treating stage 1 rhodesiense disease [102]. It is usually confined to use in Eastern and Southern Africa because in West Africa coinfection with the nematode *Onchocerca volvulus* is common and suramin's activity against microfilarial worms can induce a profound, even fatal, anaphylaxis. Typically, the drug is administered in a course of five injections, every 3–7 days over 4 weeks, with a loading dose given to ensure that patients can tolerate the drug [94]. Most drug (99%) is bound to serum protein, and the clearance half-life is remarkably long (41–78 days). Suramin fails to penetrate the blood–brain barrier, so it is useful only in stage 1. Uptake into trypanosomes involves binding of drug (as a drug-protein complex) to the ISG75 receptor on the cell surface, followed by uptake inside endocytic vesicles [103]. A high-throughput approach, to determine genes whose loss of function could lead to drug resistance, indicated that loss of ISG75 and various components of the endocytic pathway could all promote resistance. Precisely how suramin finally kills cells is not known. Although suramin resistance has rarely been reported in the field, it can be selected in the laboratory with ease, via the molecular RITseq approach and also through selecting in increasing doses [104]. Various side effects are common, including reversible nephrotoxicity, neuropathy, nausea, anemia, and urticaria [102]. When the drug was trialed in HIV patients in higher doses, these toxic effects became more pronounced; generally, however, the dosing in HAT treatment is less aggressive and thus tolerated.

39.7.3 Stage 2

39.7.3.1 Melarsoprol

Melarsoprol, a melaminophenyl-based organic arsenical, is extraordinarily toxic, with 5–10% of patients taking it experiencing a reactive encephalopathy that is fatal in around half of these cases [94]. The cause of this reactive encephalopathy is not certain. Interestingly, melarsoprol, or its *in vivo* metabolite melarsen oxide, enters mammalian cells via a thiamine transporter [105], and it would be of interest to see if heterogeneity in host thiamine transport is related to marlsoprol toxicity and, indeed, to variability in efficacy, should these transporters be involved in melarsoprol uptake in the brain. Moreover, if uptake via thiamine transporters contributes to the drug's toxicity, it might also be related to dietary status with respect to thiamine. To date, this theory is untested. Melarsoprol is given as a 3.6% solution in propylene glycol by intravenous injection, predominantly using the ten-day course [106]. This protocol replaced earlier regimens when new pharmacokinetic data revealed that the drug does not accumulate in tissues, as had been true for previous arsenic-based trypanocides.

Melarsoprol's trypanocidal mode of action has yet to be established. A variety of enzyme systems are known to be inhibited, and interactions with intracellular thiols are potent. An increased incidence of treatment failures through the 1990s led to speculation that resistance might have arisen. Resistance was shown to relate to loss of a carrier protein—the so-called P2 amino-purine transporter—that is capable of accumulating melarsen oxide into trypanosomes. [107]. The RITseq approach showed that AQP2 is also involved in resistance [103], which explains earlier work showing the HAPT1 transporter's role in melarsen oxide resistance [108], given that AQP2 is equivalent to HAPT1.

The spread of resistance and the acute toxicity associated with melarsoprol, coupled with the improved availability of eflornithine and NECT, mean that melarsoprol's use in gambiense disease is waning. An oral formulation of melarsoprol complexed with cyclodextrins [109] has been proposed as a potentially less toxic administration, particularly for rhodesiense patients for whom treatment with NECT is not currently indicated. However, it is not certain for how long melarsoprol will be produced. Because it is currently used for fewer than one thousand cases per year, the cost of production and the environmental problems associated with toxic residue from its manufacture make it hard to justify continued production. This is particularly true if NECT can find utility in the treatment of rhodesiense disease, and if compounds in the pipeline for HAT reach the clinic and can also be shown to be effective in rhodesiense disease. Notwithstanding, the risk that these newer developments will not reach the clinic means that sustained manufacture and distribution of melarsoprol should be encouraged until new options are proven and available for use.

Melarsoprol is rapidly metabolized to melarsen oxide *in vivo*, and this active compound shows a mean elimination half-life of 3.5 hours [94]. The active principle does cross the blood–brain barrier but only to levels that are around 2% of maximum plasma levels. This is sufficient to kill parasites in the CSF; however, the drug remains at trypanocidal concentrations only just long enough to effect cure. Given the fact that sensitivity levels can change through reduced expression of the P2 and/or the AQP2 transporters, risks of treatment failure are high. It has been proposed that this proximity between drug dosing into the brain and trypanocidal activity is responsible for treatment failures [110]. However, one study did show that the prevalence of alleles of the *TbAT1* gene known to be associated with treatment failure increased when treatment failures were common in northern Uganda and, then, dropped once melarsoprol had been replaced by eflornithine [111] as the primary treatment. Furthermore, heterogeneity in expression of human transporters, such as the thiamine transporter believed to play a role in uptake into mammalian cells, might contribute to host sensitivity to the drug.

Adverse effects with melarsoprol extend beyond the reactive encephalopathy with other neurological sequelae: pyrexia, headache, pruritus, and thrombocytopenia are all common, as are occasional reports of heart failure. The use of corticosteroids (prednisolone) has been reported to offer some protection against reactive encephalopathy [94].

39.7.3.2 Eflornithine

Eflornithine (D,L-α-difluoromethylornithine, marketed as Ornidyl), is an analog of the nonproteinogenic amino acid ornithine. It inhibits the enzyme ornithine decarboxylase (ODC), which is a key component of the polyamine biosynthetic pathway [112]. The drug is active against *T. b. gambiense* but less so against *T. b. rhodesiense*. The reasons for this are not known, although it has been proposed that they might relate to the fact that the target enzyme, ornithine decarboxylase, is turned over more rapidly in rhodesiense parasites than in gambiense parasites [113]; moreover, since the drug is an irreversible inhibitor of the enzyme, its relatively fast replenishment in rhodesiense could render the parasites less susceptible. The same mechanism describes the difference in susceptibility of mammalian and trypanosomal cells.

Other possible mechanisms have also been proposed [114]. For example, a relatively high uptake of putrescine, the target enzyme's product, in rhodesiense as compared to gambinese was suggested, to allow some strains of the former parasite to resist the drug, since acquisition of host-derived putrescine enables them to bypass the drug-induced lesion.

As monotherapy, eflornithine administration involves slow intravenous infusion (up to 30 minutes) of 100 mg kg^{-1} body weight drug at 6-hour intervals (i.e., 400 mg kg^{-1} per day) for 14 days.

In vitro, IC_{50} growth inhibitory values for eflornithine are high (81–693 μM). The drug's activity *in vivo* appears to depend on a functioning immune system. It has been shown that eflornithine enters trypanosomes via an amino acid transporter, TbAAT6, and resistance is easily selected in the laboratory when that transporter is lost [115].

The mean half-life in plasma following intravenous injection of eflornithine is only in the order of 3 hours. CSF-to-plasma ratios in humans have been reported between 0.1 and 0.9, but in mice apparently very little eflornithine enters the brain. It is possible that brain transporters for eflornithine exist and that differences between such transporters in mice and man could account for different PK. Polymorphisms in such a transporter could also explain variability in brain permeation of the drug and why efficacy never reaches 100% in treatment within different foci. Although generally considered safe, eflornithine's side effects have been noted. These include fever, headache, hypertension, macular rash, peripheral neuropathy, tremor, and gastrointestinal problems such as diarrhea.

39.7.3.3 Nifurtimox-Eflornithine Combination Therapy

Antimicrobial combination chemotherapies are increasingly used because they improve efficacy, decrease dosing, and reduce the risk of drug resistance. In a series of clinical trials involving various combinations of registered trypanocides with nifurtimox (previously used exclusively for Chagas disease), it was concluded that eflornithine with nifurtimox was the most successful treatment [116]. A regimen using eflornithine by intravenous infusion at 200 mg/kg every 12 hours for 7 days (rather than every 6 hours for 14 days, as in monotherapy), with nifurtimox given orally three times a day for 10 days, offers significant advantages in cost and convenience yet is of equal efficacy to that of the longer eflornithine monotherapy.

Curiously, eflornithine and nifurtimox are not synergistic in *in vitro* assays [117], possibly because, rather than oxidative stress as previously assumed, nifurtimox's mode of action appears to involve reduction by an unusual nitroreductase in *T. brucei* followed by further metabolism to a highly reactive nitrile derivative. Loss of the nitroreductase causes resistance to nitroheterocycles [118]. Clearance of nifurtimox is fast, with a plasma elimination half-life of approximately 3 hours, but the drug can accumulate across the blood–brain barrier to levels approximately half of those found in plasma [94]. African trypanosomes are not very susceptible to nifurtimox, with IC_{50} values of approximately $5\,\mu M$ *in vitro*, and resistance is easily selected in the laboratory. Toxic effects to the central and peripheral nervous systems have been reported with nifurtimox.

In the combination therapy trial, the only significant adverse events reported to occur at a higher frequency in the nifurtimox-eflornithine combination therapy cohort versus the eflornithine monotherapy cohort were an increased incidence of tremors and increased gastrointestinal disturbance, including anorexia and nausea. NECT has thus become the treatment of choice for stage 2 gambiense disease.

39.7.4 Compounds in Clinical or Advanced Preclinical Trials

39.7.4.1 Fexinidazole

Fexinidazole, a 2-substituted 5-nitroimidazole, was originally synthesized by Hoechst in the 1970s and shown to have trypanocidal activity [119]. Clinical development started nearly 40 years later, when the Drugs for Neglected Diseases initiative (DNDi) investigated the potential of nitroheterocyclic compounds as potential drug candidates for treatment of stage 2 HAT [120]. In spite of the feared genotoxicity potential and expected Ames test positivity, fexinidazole was found to be safe in phase I trials, where its metabolism to sulfoxide and sulfone derivatives was confirmed. These metabolites have trypanocidal effect similar to that of the parent compound. Although substantially lower than that of other trypanocides, such as melarsoprol, suramin, and pentamidine, this effect remains at trypanocidal levels in the brain long enough to cure a stage 2 infection [120].

Fexinidazole succeeded in phase I trials, although substantial food effects were noted, and it is currently undergoing evaluation in a phase II/III trial. If successful, it will be the first orally available drug suitable for use in stage 2 HAT and promises to revolutionize treatment by removing the need for parenteral administration.

39.7.4.2 SCYX-7158

A second compound, SCYX-7158, recently entered phase I clinical trials following successful progression through preclinical testing [121]. The benzoxaborole-6-carboxamide class, initially developed by Anacor, was expanded by Scynexis, which combined *in vitro* potency data against trypanosomes with *in vivo* pharmacokinetic data in mice to work on structures whose brain concentrations stayed high enough, for long enough, to effectively cure stage 2 infections [122]. The compound has a long terminal half life in humans, and phase I safety testing is under way, with entry into phase II trials planned for 2014 if the safety profile is satisfactory.

39.7.4.3 Dicationic Molecules

A third class of compounds that has shown activity in stage 2 disease in rodents and primates comprises the aza-diamidines. These molecules are analogs of DB75 (furamidine) and are the active principle of the prodrug DB289 (pafuramidine). Pafuramidine was dropped from clinical testing following the identification of unacceptable nephrotoxicity in an extended phase I safety trial [100], in spite of its having proven efficacious and relatively safe in a phase III trial. CPD-0802 (or DB829 as the hydrochloride salt) and its prodrug, DB868, are aza analogs of DB75 and DB289, in both of which the ring structure possesses a nitrogen. The active diamidine compounds are highly potent in rodent and primate models (123). The corresponding asymmetric compounds, in which just a single ring contains the nitrogen, demonstrate similar potency. Further work is underway to assess the toxic potential of these compounds compared with that of DB75, given the unexpected nephrotoxicity that led to the cessation of pafuramidine development. Current information suggests that the aza diamidines have a greatly reduced propensity to accumulate in the kidney, and there is hope that these compounds can be considered for entry into clinical trials at some point.

39.8 CONCLUSIONS AND FUTURE PERSPECTIVES

Human African trypanosomiasis is caused by protozoa of the *Trypanosoma brucei* subgroup. Classically, the disease is stratified based on its causative subspecies (*T. b. gambiense* versus *T. b. rhodesiense*) and whether it is in stage 1 (hemolymphatic) or stage 2 (neurological). Recently a third type of infection was discovered in which *T. evansi* can infect humans defective in specific genes that confer resistance to most trypanosome species.

Trypanosome lytic factors, including apolipoprotein L1 (APOL1), have been selected to combat trypanosome infections. However, the human infectious subspecies have evolved mechanisms to counter APOL1 activity. *T. evansi* infections have been seen that involve rare genetic defects in *APOL1*. Other genotypes (e.g., some cytokine gene alleles) are also associated with the severity of trypanosome infection. Pathogenesis depends on parasite and host genes in addition to those genes already characterized.

Recent successes in HAT control might mean that there will be insufficient human cases to provide the statistical power to find other genetic associations to the disease status. However, caution must be exercised in considering HAT to be a disease whose decline means that further efforts at control are unnecessary. It is a disease with the capacity for rapid resurgence; thus, enhanced efforts toward its elimination are required. Surveillance should be maintained and efforts to improve diagnostic capability and chemotherapy continued.

GLOSSARY

HAT (sleeping sickness) Human African trypanosomiasis is the correct term for human infections caused by the African trypanosomes of the *T. brucei* species: *Trypanosoma brucei gambiense* and *T. b. rhodesiense*. The disease is commonly referred to as sleeping sickness when in its second, CNS involved stage. At this stage, trypanosomes cause various perturbations to the neurological system, including the changes to sleep/wake patterns that led to the disease acquiring its common name

Serum resistance–associated (*SRA*) gene The *SRA* is found in *Trypanosoma brucei rhodesiense*. It is derived from a *VSG* gene and is expressed from a single expression site; hence, trypanosomes carrying the gene are only resistant to serum when the particular *VSG* expression site with which they are associated is expressed. Expression sites are found at the telomeres of the parasite's chromosomes and contain a strong RNA polymerase 1 promoter that allows expression of a single *VSG* gene at a time, in addition to several expression site–associated genes (*ESAGS*) that are also found there

***T. b. gambiense*–specific glycoprotein gene (*TGSGP*)** *TGSGP*, like *SRA* in *T. b. rhodesiense*, is found only in one subspecies. It is also a mutated version of a variant surface glycoprotein gene and can confer resistance against trypanosome lytic factors in human serum, apparently by influencing membrane fluidity in ways that affect uptake of TLF

Trypanosome lytic factors (TLFs) TLFs are components of human blood that are capable of killing trypanosomes. Initially identified as being part of the high-density lipoprotein fraction of plasma, it was shown that a key individual component was apolipoprotein L1 (APOL1). APOL1 is capable of lysing trypanosomes other than those subspecies that have evolved molecules that neutralize its activity. Other components—such as a haptoglobin-like protein that is part of the complex particle that also contains APOL1—are required for the binding of these particles to the trypanosome and contribute to activity since specific receptors are required for this binding

Tsetse fly The tsetse fly is a dipteran insect of the genus *Glossina*. Various species can transmit trypanosomes, which replicate within the tsetse fly midgut as procyclic forms. Some of these forms migrate to the salivary gland, where they transform to metacyclic forms that are preadapted to survive within the insect host

Variant surface glycoproteins (VSGs) VSGs are expressed one at a time, and acquire a glycosylphosphatidyl inositol (GPI) anchor that links them to the plasma membrane. The trypanosome has up to a thousand different *VSG* genes, but expresses only one of them at a time. The proteins form dimers that pack tightly at the parasite's surface and prevent host immunological proteins, including components of the complement complex and immunoglobulins, from reaching the cell surface. Eventually the host produces antibodies that bind to a given antigen type, and cells carrying that VSG are killed. However, a small number of parasites change expression to a new VSG that is not recognized by those antibodies. These cells then proliferate until a specific antibody response is raised to them. In this way, the parasites can evade the immune system *ad infinitum*

REFERENCES

[1] Brun R, Blum J, Chappuis F, Burri C. Human African trypanosomiasis. Lancet 2010;375(9709):148–59. 9
[2] Barrett MP, Burchmore RJ, Stich A, Lazzari JO, Frasch AC, Cazzulo JJ, et al. The trypanosomiases. Lancet 2003;362(9394):1469–80.
[3] Steverding D. The history of African trypanosomiasis. Parasit Vectors 2008;1(1):3.
[4] Chappuis F, Loutan L, Simarro P, Lejon V, Buscher P. Options for field diagnosis of human African trypanosomiasis. Clin Microbiol Rev 2005;18(1):133–46.
[5] Vanhollebeke B, Truc P, Poelvoorde P, Pays A, Joshi PP, Katti R, et al. Human trypanosoma evansi infection linked to a lack of apolipoprotein L-I. N Engl J Med 2006;355(26):2752–6.
[6] Picozzi K, Fevre EM, Odiit M, Carrington M, Eisler MC, Maudlin I, et al. Sleeping sickness in Uganda: a thin line between two fatal diseases. BMJ 2005;331(7527):1238–41.
[7] Simarro PP, Franco J, Diarra A, Postigo JA, Jannin J. Update on field use of the available drugs for the chemotherapy of human African trypanosomiasis. Parasitology 2012;139(7):842–6.
[8] Fevre EM, Coleman PG, Welburn SC, Maudlin I. Reanalyzing the 1900–1920 sleeping sickness epidemic in Uganda. Emerg Infect Dis 2004;10(4):567–73.
[9] Aksoy S, Gibson WC, Lehane MJ. Interactions between tsetse and trypanosomes with implications for the control of trypanosomiasis. Adv Parasitol 2003;53:1–83.
[10] Njiokou F, Laveissiere C, Simo G, Nkinin S, Grebaut P, Cuny G, et al. Wild fauna as a probable animal reservoir for *Trypanosoma brucei gambiense* in Cameroon. Infect Genet Evol 2006;6(2):147–53.
[11] Simarro PP, Cecchi G, Franco JR, Paone M, Diarra A, Ruiz-Postigo JA, et al. Estimating and mapping the population at risk of sleeping sickness. PLoS Negl Trop Dis 2012;6(10):e1859.
[12] Louis FJ, Somarro PP. Les difficiles débuts de la lutte contre la maladie du sommeil en Afrique équatorielle française [Difficult start for the campaign against sleeping sickness in French equatorial Africa]. Médecine Tropicale 2005;65:251–7.
[13] Maurice J. New WHO plan targets the demise of sleeping sickness. Lancet 2013;381(9860):13–14.
[14] Rassi Jr A, Rassi A, Marin-Neto JA. Chagas disease. Lancet 2010;375(9723):1388–402.
[15] De GC, Hamers R. The serum resistance-associated (SRA) gene of *Trypanosoma brucei rhodesiense* encodes a variant surface glycoprotein-like protein. Mol Biochem Parasitol 1994;68(2):277–84.
[16] Pays E, Vanhollebeke B. Mutual self-defence: the trypanolytic factor story. Microbes Infect 2008;10(9):985–9.
[17] Berberof M, Perez-Morga D, Pays E. A receptor-like flagellar pocket glycoprotein specific to *Trypanosoma brucei gambiense*. Mol Biochem Parasitol 2001;113(1):127–38.
[18] Bromidge T, Gibson W, Hudson K, Dukes P. Identification of *Trypanosoma brucei gambiense* by PCR amplification of variant surface glycoprotein genes. Acta Trop 1993;53(2):107–19.
[19] Rifkin MR. Identification of the trypanocidal factor in normal human serum: high density lipoprotein. Proc Natl Acad Sci USA 1978;75(7):3450–4.

[20] Wheeler RJ. The trypanolytic factor-mechanism, impacts and applications. Trends Parasitol 2010;26(9):457–64.

[21] Vanhollebeke B, De MG, Nielsen MJ, Pays A, Tebabi P, Dieu M, et al. A haptoglobin-hemoglobin receptor conveys innate immunity to *Trypanosoma brucei* in humans. Science 2008;320(5876):677–81.

[22] Lecordier L, Vanhollebeke B, Poelvoorde P, Tebabi P, Paturiaux-Hanocq F, Andris F, et al. C-terminal mutants of apolipoprotein L-I efficiently kill both *Trypanosoma brucei brucei* and *Trypanosoma brucei rhodesiense*. PLoS Pathog 2009;5(12):e1000685.

[23] Gibson WC. The SRA gene: the key to understanding the nature of *Trypanosoma brucei rhodesiense*. Parasitology 2005;131(Pt 2):143–50.

[24] Kieft R, Capewell P, Turner CM, Veitch NJ, Macleod A, Hajduk S. Mechanism of *Trypanosoma brucei gambiense* (group 1) resistance to human trypanosome lytic factor. Proc Natl Acad Sci USA 2010;107(37):16137–41.

[25] Joshi PP, Chaudhari A, Shegokar VR, Powar RM, Dani VS, Somalwar AM, et al. Treatment and follow-up of the first case of human trypanosomiasis caused by *Trypanosoma evansi* in India. Trans R Soc Trop Med Hyg 2006;100(10):989–91.

[26] Brun R, Hecker H, Lun ZR. *Trypanosoma evansi* and *T. equiperdum*: distribution, biology, treatment and phylogenetic relationship (a review). Vet Parasitol 1998 October;79(2):95–107.

[27] Lamour N, Riviere L, Coustou V, Coombs GH, Barrett MP, Bringaud F. Proline metabolism in procyclic *Trypanosoma brucei* is down-regulated in the presence of glucose. J Biol Chem 2005;280(12):11902–10.

[28] Besteiro S, Barrett MP, Riviere L, Bringaud F. Energy generation in insect stages of *Trypanosoma brucei*: metabolism in flux. Trends Parasitol 2005;21(4):185–91.

[29] Schnaufer A, Domingo GJ, Stuart K. Natural and induced dyskinetoplastic trypanosomatids: how to live without mitochondrial DNA. Int J Parasitol 2002;32(9):1071–84.

[30] Wasser WG, Tzur S, Wolday D, Adu D, Baumstein D, Rosset S, et al. Population genetics of chronic kidney disease: the evolving story of APOL1. J Nephrol 2012;25(5):603–18.

[31] Freedman BI, Langefeld CD, Turner J, Nunez M, High KP, Spainhour M, et al. Association of APOL1 variants with mild kidney disease in the first-degree relatives of African American patients with non-diabetic end-stage renal disease. Kidney Int 2012;82(7):805–11.

[32] Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI, et al. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. Science 2010;329(5993):841–5.

[33] Lopez C, Saravia C, Gomez A, Hoebeke J, Patarroyo MA. Mechanisms of genetically-based resistance to malaria. Gene 2010;467(1–2):1–12.

[34] Bunn H. The triumph of good over evil: protection by the sickle gene against malaria. Blood 2013;121(1):20–5.

[35] Beutler E. G6PD: population genetics and clinical manifestations. Blood Rev 1996;10(1):45–52.

[36] Howes RE, Battle KE, Satyagraha AW, Baird J, Hay SI. G6PD deficiency: global distribution. Genet Variants Primaquine Ther 2013;81:133–201.

[37] Maclean L, Chisi JE, Odiit M, Gibson WC, Ferris V, Picozzi K, et al. Severity of human african trypanosomiasis in East Africa is associated with geographic location, parasite genotype, and host inflammatory cytokine response profile. Infect Immun 2004;72(12):7040–4.

[38] Sternberg JM, Maclean L. A spectrum of disease in human African trypanosomiasis: the host and parasite genetics of virulence. Parasitology 2010;137(14):2007–15.

[39] Bucheton B, Macleod A, Jamonneau V. Human host determinants influencing the outcome of *Trypanosoma brucei gambiense* infections. Parasite Immunol 2011;33(8):438–47.

[40] Jamonneau V, Ravel S, Garcia A, Koffi M, Truc P, Laveissiere C, et al. Characterization of *Trypanosoma brucei* s.l. infecting asymptomatic sleeping-sickness patients in Cote d'Ivoire: a new genetic group? Ann Trop Med Parasitol 2004;98(4):329–37.

[41] Garcia A, Jamonneau V, Magnus E, Laveissiere C, Lejon V, N'Guessan P, et al. Follow-up of card agglutination trypanosomiasis test (CATT) positive but apparently aparasitaemic individuals in Cote d'Ivoire: evidence for a complex and heterogeneous population. Trop Med Int Health 2000;5(11):786–93.

[42] Holzmuller P, Biron DG, Courtois P, Koffi M, Bras-Goncalves R, Daulouede S, et al. Virulence and pathogenicity patterns of *Trypanosoma brucei gambiense* field isolates in experimentally infected mouse: differences in host immune response modulation by secretome and proteomics. Microbes Infect 2008;10(1):79–86.

[43] Giroud C, Ottones F, Coustou V, Dacheux D, Biteau N, Miezan B, et al. Murine models for *Trypanosoma brucei gambiense* disease progression-from silent to chronic infections and early brain tropism. PLoS Negl Trop Dis 2009;3(9):e509.

[44] Drennan MB, Stijlemans B, Van Den Abbeele J, Quesniaux VJ, Barkhuizen M, Brombacher F, et al. The induction of a type 1 immune response following a *Trypanosoma brucei* infection is MyD88 dependent. J Immunol 2005;175(4):2501–9.

[45] Duleu S, Vincendeau P, Courtois P, Semballa S, Lagroye I, Daulouede S, et al. Mouse strain susceptibility to trypanosome infection: an arginase-dependent effect. J Immunol 2004;172(10):6298–303.

[46] Lejon V, Lardon J, Kenis G, Pinoges L, Legros D, Bisser S, et al. Interleukin (IL)-6, IL-8 and IL-10 in serum and CSF of *Trypanosoma brucei gambiense* sleeping sickness patients before and after treatment. Trans Roy Soc Trop Med Hyg 2002;96(3):329–33.

[47] Iraqi F, Clapcott SJ, Kumari P, Haley CS, Kemp SJ, Teale AJ. Fine mapping of trypanosomiasis resistance loci in murine advanced intercross lines. Mamm Genome 2000;11(8):645–8.

[48] Clapcott SJ, Teale AJ, Kemp SJ. Evidence for genomic imprinting of the major QTL controlling susceptibility to trypanosomiasis in mice. Parasite Immunol 2000;22(5):259–63.

[49] Courtin D, Argiro L, Jamonneau V, N'Dri L, N'Guessan P, Abel L, et al. Interest of tumor necrosis factor-alpha -308 G/A and interleukin-10 -592 C/A polymorphisms in human African trypanosomiasis. Infect Genet Evol 2006;6(2):123–9.

[50] Stijlemans B, Guilliams M, Raes G, Beschin A, Magez S, De BP. African trypanosomiasis: from immune escape and immunopathology to immune intervention. Vet Parasitol 2007;148(1):3–13.

[51] Horn D, McCulloch R. Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Curr Opin Microbiol 2010;13(6):700–5.

[52] Barry JD, McCulloch R. Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. Adv Parasitol 2001;49:1–70.

[53] Morrison LJ, Marcello L, McCulloch R. Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. Cell Microbiol 2009;11(12):1724–34.

[54] Darji A, Lucas R, Magez S, Torreele E, Palacios J, Sileghem M, et al. Mechanisms underlying trypanosome-elicited immunosuppression. Ann Soc Belg Med Trop 1992;72(Suppl. 1):27–38.

[55] Sternberg JM. Immunobiology of African trypanosomiasis. Chem Immunol 1998;70:186–99.

[56] Gomez-Rodriguez J, Stijlemans B, De MG, Korf H, Brys L, Berberof M, et al. Identification of a parasitic immunomodulatory protein triggering the development of suppressive M1 macrophages during African trypanosomiasis. J Infect Dis 2009;200(12):1849–60.

[57] Checchi F, Filipe JA, Barrett MP, Chandramohan D. The natural progression of Gambiense sleeping sickness: what is the evidence? PLoS Negl Trop Dis 2008;2(12):e303.

[58] Poltera AA. Pathology of human African trypanosomiasis with reference to experimental African trypanosomiasis and infections of the central nervous system. Br Med Bull 1985;41(2):169–74.

[59] Kennedy PGE. Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). Lancet Neurol 2013;12(2):186–94.

[60] Scott JA, Davidson RN, Moody AH, Bryceson AD. Diagnosing multiple parasitic infections: trypanosomiasis, loiasis and schistosomiasis in a single case. Scand J Infect Dis 1991;23(6):777–80.

[61] Blum JA, Burri C, Hatz C, Kazumba L, Mangoni P, Zellweger MJ. Sleeping hearts: the role of the heart in sleeping sickness (human African trypanosomiasis). Trop Med Int Health 2007;12(12):1422–32.

[62] Buguet A, Gati R, Sevre JP, Develoux M, Bogui P, Lonsdorfer J. 24 hour polysomnographic evaluation in a patient with sleeping sickness. Electroencephalogr Clin Neurophysiol 1989;72(6):471–8.

[63] Mpandzou G, Cespuglio R, Ngampo S, Bandzouzi B, Bouteille B, Vincendeau P, et al. Polysomnography as a diagnosis and post-treatment follow-up tool in human African trypanosomiasis: a case study in an infant. J Neurol Sci 2011;305(1–2):112–5.

[64] Reincke M, Arlt W, Heppner C, Petzke F, Chrousos GP, Allolio B. Neuroendocrine dysfunction in African trypanosomiasis. The role of cytokines. Ann N Y Acad Sci 1998;840:809–21.

[65] Ikede BO, Elhassan E, Akpavie SO. Reproductive disorders in African trypanosomiasis: a review. Acta Trop 1988;45(1):5–10.

[66] Pepin J, Ethier L, Kazadi C, Milord F, Ryder R. The impact of human immunodeficiency virus infection on the epidemiology and treatment of *Trypanosoma brucei gambiense* sleeping sickness in Nioki, Zaire. Am J Trop Med Hyg 1992;47(2):133–40.

[67] Blum J, Nkunku S, Burri C. Clinical description of encephalopathic syndromes and risk factors for their occurrence and outcome during melarsoprol treatment of human African trypanosomiasis. Trop Med Int Health 2001;6(5):390–400.

[68] Blum JA, Neumayr AL, Hatz CF. Human African trypanosomiasis in endemic populations and travellers. Eur J Clin Microbiol Infect Dis 2012;31(6):905–13.

[69] Magnus E, Vervoort T, Van MN. A card-agglutination test with stained trypanosomes (C.A.T.T.) for the serological diagnosis of *T. B. gambiense* trypanosomiasis. Ann Soc Belg Med Trop 1978;58(3):169–76.

[70] Buscher P, Gilleman Q, Lejon V. Rapid diagnostic test for sleeping sickness. N Engl J Med 2013 March 14;368(11):1069–70.

[71] Dukes P, Gibson WC, Gashumba JK, Hudson KM, Bromidge TJ, Kaukus A, et al. Absence of the LiTat 1.3 (CATT antigen) gene in *Trypanosoma brucei gambiense* stocks from Cameroon. Acta Trop 1992;51(2):123–34.

[72] Magnus E, Van MN, Vervoort T, Le RD, Wery M. Use of freeze-dried trypanosomes in the indirect fluorescent antibody test for the serodiagnosis of sleeping sickness. Ann Soc Belg Med Trop 1978;58(2):103–9.

[73] Vervoort T, Magnus E, Van MN. Enzyme-linked immunosorbent assay (ELISA) with variable antigen for serodiagnosis of *T. B. gambiense* trypanosomiasis. Ann Soc Belg Med Trop 1978;58(3):177–83.

[74] Jones D, Nyalwidhe J, Tetley L, Barrett MP. McArthur revisited: fluorescence microscopes for field diagnostics. Trends Parasitol 2007;23(10):468–9.

[75] Bieler S, Matovu E, Mitashi P, Ssewannyana E, Bi Shamamba SK, Bessell PR, et al. Improved detection of *Trypanosoma brucei* by lysis of red blood cells, concentration and LED fluorescence microscopy. Acta Trop 2012;121(2):135–40.

[76] Woo PT. The haematocrit centrifuge technique for the diagnosis of African trypanosomiasis. Acta Trop 1970;27(4):384–6.

[77] Bailey JW, Smith DH. The use of the acridine orange QBC technique in the diagnosis of African trypanosomiasis. Trans R Soc Trop Med Hyg 1992;86(6):630.

[78] Lumsden WH, Kimber CD, Evans DA, Doig SJ. *Trypanosoma brucei*: miniature anion-exchange centrifugation technique for detection of low parasitaemias: adaptation for field use. Trans R Soc Trop Med Hyg 1979;73(3):312–7.

[79] Solano P, Jamonneau V, N'Guessan P, N'Dri L, Dje NN, Miezan TW, et al. Comparison of different DNA preparation protocols for PCR diagnosis of human African trypanosomosis in Cote d'Ivoire. Acta Trop 2002;82(3):349–56.

[80] Deborggraeve S, Buscher P. Recent progress in molecular diagnosis of sleeping sickness. Expert Rev Mol Diagn 2012;12(7):719–30.

[81] Kuboki N, Inoue N, Sakurai T, Di CF, Grab DJ, Suzuki H, et al. Loop-mediated isothermal amplification for detection of African trypanosomes. J Clin Microbiol 2003;41(12):5517–24.

[82] Kennedy PG. Diagnosing central nervous system trypanosomiasis: two stage or not to stage? Trans R Soc Trop Med Hyg 2008;102(4):306–7.

[83] Doua F, Miezan TW, Sanon Singaro JR, Boa YF, Baltz T. The efficacy of pentamidine in the treatment of early-late stage *Trypanosoma brucei gambiense* trypanosomiasis. Am J Trop Med Hyg 1996;55(6):586–8.

[84] Lejon V, Buscher P. Diagnosis of the disease stage in sleeping sickness: towards a new approach. Bulletin de la Societe de Pathologie Exotique 2002;95(5):338–40.

[85] Lejon V, Buscher P, Sema NH, Magnus E, Van MN. Human African trypanosomiasis: a latex agglutination field test for quantifying IgM in cerebrospinal fluid. Bull World Health Organ 1998;76(6):553–8.

[86] Hainard A, Tiberti N, Robin X, Lejon V, Ngoyi DM, Matovu E, et al. A combined CXCL10, CXCL8 and H-FABP panel for the staging of human African trypanosomiasis patients. PLoS Negl Trop Dis 2009;3(6):e459.

[87] Tiberti N, Matovu E, Hainard A, Enyaru JC, Lejon V, Robin X, et al. New biomarkers for stage determination in *Trypanosoma brucei rhodesiense* sleeping sickness patients. Clin Transl Med 2013;2(1):1.

[88] Hainard A, Tiberti N, Robin X, Ngoyi DM, Matovu E, Enyaru JC, et al. Matrix metalloproteinase-9 and intercellular adhesion molecule 1 are powerful staging markers for human African trypanosomiasis. Trop Med Int Health 2011;16(1):119–26.

[89] Tiberti N, Hainard A, Lejon V, Courtioux B, Matovu E, Enyaru C, et al. Cerebrospinal fluid neopterin as marker of the meningo-encephalitic stage of *Trypanosoma brucei gambiense* sleeping sickness. Plos One 2012;7(7).

[90] Hall JE, Seed JR. Increased urinary excretion of aromatic amino acid catabolites by *Microtus montanus* chronically infected with *Trypanosoma brucei gambiense*. Comp Biochem Physiol B 1984;77(4):755–60.

[91] Hall JE, Seed JR, Sechelski JB. Multiple alpha-keto aciduria in *Microtus montanus* chronically infected with *Trypanosoma brucei gambiense*. Comp Biochem Physiol B 1985;82(1):73–8.

[92] Wang Y, Utzinger J, Saric J, Li JV, Burckhardt J, Dirnhofer S, et al. Global metabolic responses of mice to *Trypanosoma brucei brucei* infection. Proc Natl Acad Sci USA 2008;105(16):6127–32.

[93] Creek DJ, Anderson J, McConville MJ, Barrett MP. Metabolomic analysis of trypanosomatid protozoa. Mol Biochem Parasitol 2012;181(2):73–84.

[94] Barrett M, Boykin D, Brun R, Tidwell R. Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. British J Pharmacol 2007;152(8):1155–71.

[95] Sands M, Kron MA, Brown RB. Pentamidine: a review. Rev Infect Dis 1985;7(5):625–34.

[96] Creek DJ, Nijagal B, Kim DH, Rojas F, Matthews KR, Barrett MP. Metabolomics guides rational development of a simplified cell culture medium for drug screening against *Trypanosoma brucei*. Antimicrob Agents Chemother 2013;57(6):2768–79.

[97] Baker N, Glover L, Munday JC, Aguinaga AD, Barrett MP, de Koning HP, et al. Aquaglyceroporin 2 controls susceptibility to melarsoprol and pentamidine in African trypanosomes. Proc Natl Acad Sci USA 2012;109(27):10996–1001. 3

[98] de Koning HP. Uptake of pentamidine in *Trypanosoma brucei brucei* is mediated by three distinct transporters: implications for cross-resistance with arsenicals. Mol Pharmacol 2001;59(3):586–92.

[99] Creek DJ, Barrett MP. Determination of antiprotozoal drug mechanisms by metabolomics approaches. Parasitology 2013;5:1–10.

[100] Paine MF, Wang MZ, Generaux CN, Boykin DW, Wilson WD, de Koning HP, et al. Diamidines for human African trypanosomiasis. Curr Opin Investig Drugs 2010;11(8):876–83.

[101] Ming X, Ju W, Wu H, Tidwell RR, Hall JE, Thakker DR. Transport of dicationic drugs pentamidine and furamidine by human organic cation transporters. Drug Metab Dispos 2009;37(2):424–30.

[102] Voogd TE, Vansterkenburg EL, Wilting J, Janssen LH. Recent research on the biological activity of suramin. Pharmacol Rev 1993;45(2):177–203.

[103] Alsford S. High-throughput decoding of antitrypanosomal drug efficacy and resistance 2012;482(7384):232–36.

[104] Scott AG, Tait A, Turner CM. Characterisation of cloned lines of *Trypanosoma brucei* expressing stable resistance to MelCy and suramin. Acta Trop 1996;60(4):251–62.

[105] Szyniarowski P, Bettendorff L, Schweingruber ME. The antitrypanosomal drug melarsoprol competitively inhibits thiamin uptake in mouse neuroblastoma cells. Cell Biol Toxicol 2006;22(3):183–7.

[106] Burri C, Nkunku S, Merolle A, Smith T, Blum J, Brun R. Efficacy of new, concise schedule for melarsoprol in treatment of sleeping sickness caused by *Trypanosoma brucei gambiense*: a randomised trial. Lancet 2000;355(9213):1419–25. 22

[107] Carter NS, Fairlamb AH. Arsenical-resistant trypanosomes lack an unusual adenosine transporter. Nature 1993;361(6408):173–6.

[108] Bridges DJ, Gould MK, Nerima B, Maser P, Burchmore RJ, de Koning HP. Loss of the high-affinity pentamidine transporter is responsible for high levels of cross-resistance between arsenical and diamidine drugs in African trypanosomes. Mol Pharmacol 2007;71(4):1098–108.

[109] Rodgers J, Jones A, Gibaud S, Bradley B, McCabe C, Barrett MP, et al. Melarsoprol cyclodextrin inclusion complexes as promising oral candidates for the treatment of human African trypanosomiasis. PLoS Negl Trop Dis 2011;5(9):e1308.

[110] Brun R, Schumacher R, Schmid C, Kunz C, Burri C. The phenomenon of treatment failures in human African trypanosomiasis. Trop Med Int Health 2001;6(11):906–14.

[111] Kazibwe AJN, Nerima B, de Koning HP, Maser P, Barrett MP, Matovu E. Genotypic status of the TbAT1/P2 adenosine transporter of *Trypanosoma brucei gambiense* isolates from Northwestern Uganda following melarsoprol withdrawal. Plos Neg Trop Dis 2009;3(9).

[112] Burri C, Brun R. Eflornithine for the treatment of human African trypanosomiasis. Parasitol Res 2003;90(Supp 1):S49–52.

[113] Iten M, Mett H, Evans A, Enyaru JC, Brun R, Kaminsky R. Alterations in ornithine decarboxylase characteristics account for tolerance of *Trypanosoma brucei rhodesiense* to D, L-alpha-difluoromethylornithine. Antimicrob Agents Chemother 1997;41(9):1922–5.

[114] Bacchi CJ, Nathan HC, Livingston T, Valladares G, Saric M, Sayer PD, et al. Differential susceptibility to DL-alpha-difluoromethylornithine in clinical isolates of *Trypanosoma brucei rhodesiense*. Antimicrob Agents Chemother 1990;34(6):1183–8.

[115] Vincent IM, Creek D, Watson DG, Kamleh MA, Woods DJ, Wong PE, et al. A molecular mechanism for eflornithine resistance in African trypanosomes. PLoS Pathog 2010;6(11):e1001204.

[116] Priotto G, Kasparian S, Mutombo W, Ngouama D, Ghorashian S, Arnold U, et al. Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. Lancet 2009;374(9683):56–64.

[117] Vincent IM, Creek DJ, Burgess K, Woods DJ, Burchmore RJ, Barrett MP. Untargeted metabolomics reveals a lack of synergy between nifurtimox and eflornithine against *Trypanosoma brucei*. PLoS Negl Trop Dis 2012;6(5):e1618.

[118] Wilkinson SR, Taylor MC, Horn D, Kelly JM, Cheeseman I. A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. Proc Natl Acad Sci USA 2008;105(13):5022–7.

[119] Jennings FW, Urquhart GM. The use of the 2 substituted 5-nitroimidazole, Fexinidazole (Hoe 239) in the treatment of chronic *T. brucei* infections in mice. Z Parasitenkd 1983;69(5):577–81.

[120] Torreele E, Bourdin TB, Tweats D, Kaiser M, Brun R, Mazue G, et al. Fexinidazole—a new oral nitroimidazole drug candidate entering clinical development for the treatment of sleeping sickness. PLoS Negl Trop Dis 2010;4(12):e923.

[121] Jacobs RT, Nare B, Wring SA, Orr MD, Chen D, Sligar JM, et al. SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human African trypanosomiasis. PLoS Negl Trop Dis 2011;5(6):e1151.

[122] Jacobs RT, Plattner JJ, Nare B, Wring SA, Chen D, Freund Y, et al. Benzoxaboroles: a new class of potential drugs for human African trypanosomiasis. Future Med Chem 2011;3(10):1259–78.

This page is intentionally left blank

Part VIII

Implementation and Regulatory

This page is intentionally left blank

Chapter 40

Implementing Clinical Pharmacogenetics: Point-of-Care and Pre-Emptive Testing

Stuart A. Scott

Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, USA

Chapter Outline

40.1 Overview 921
40.2 Clinical Pharmacogenetics and Implementation 922
40.2.1 Validity and Utility 922
40.2.2 Education, Guidelines, and Recommendations 922
40.2.3 Pharmacogenetic Testing Regulation 924
40.2.4 Pharmacogenetic Testing Turnaround Time 924
40.3 Point-of-Care Pharmacogenetic Testing 924
40.4 Pre-Emptive Pharmacogenetic Testing 925
40.4.1 Vanderbilt University Medical Center: PREDICT 926
40.4.2 University of Florida and Shands Hospital: PMP 926
40.4.3 University of Chicago: 1200 Patients Project 926
40.4.4 St. Jude Children's Research Hospital: PG4KDS 927
40.4.5 Icahn School of Medicine at Mount Sinai: CLIPMERGE–PGx 927
40.4.6 PGRN: Translational Pharmacogenetics Program 927
40.5 Conclusion and Future Perspectives 927
References 928

40.1 OVERVIEW

Clinical genetic testing began over 30 years ago with the identification of the β-globin mutation responsible for sickle cell disease [1]. As the genetic causes of many other disorders have been identified, molecular testing has grown dramatically to include population carrier-screening programs for autosomal recessive disorders [2], targeted genotyping and gene sequencing for molecular diagnosis, prenatal mutation analyses, and preimplantation genetic diagnosis. These testing scenarios predominantly involve Mendelian disorders; however, recent candidate gene and genome-wide association studies (GWASs) have identified variant alleles that contribute to some common diseases and complex traits, prompting the possibility of predictive genetic testing to evaluate personalized disease risk. Testing for these variant alleles is not common beyond direct-to-consumer (DTC) testing companies; however, as these associations become more robust, it is possible that predictive genetic testing may become more common in clinical laboratories. Importantly, early recommendations for DTC companies and the personalized medicine community stated that inclusion of pharmacogenetic variants in these testing panels was warranted as they were perceived to be more clinically actionable than the predictive markers for later-onset diseases [3].

The role of germline genetic variation in drug response has been studied since the 1950s [4], and clinical testing for selected genes known to influence drug efficacy and/or toxicity has been available for several years. However, clinical adoption of pharmacogenetic testing has remained slow despite U.S. Food and Drug Administration (FDA) product insert relabeling with pharmacogenetic information for certain drugs. A number of barriers to clinical implementation have been identified [5], including the lack of professional education and practice guidelines; concerns over clinical utility, regulatory, and reimbursement issues; and the feasibility of integrating rapid-turnaround genetic testing into routine clinical practice [6], which is often required for the variants that are currently considered actionable. To address the specific challenge of testing feasibility and turnaround time, both point-of-care and pre-emptive pharmacogenetic testing programs,

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00040-2
© 2014 Elsevier Inc. All rights reserved.

initially centered on cardiovascular pharmacogenetics, have recently been deployed. This chapter aims to review pharmacogenetic implementation with a focus on testing for germline genetic variants implicated in drug response variability and their use at the point of care.

40.2 CLINICAL PHARMACOGENETICS AND IMPLEMENTATION

Pharmacogenetic discoveries are often followed by interest in clinical implementation. However, identification of genetic markers associated with drug response does not always equate to clinically useful predictors of efficacy or adverse outcomes, and independent replication of genotype–phenotype association is essential prior to pursuing clinical implementation [7]. Several academic health centers have invested in clinical pharmacogenetics and view its implementation as a logical first step toward incorporating genomics into routine and personalized healthcare; however, a number of challenges and barriers exist for widespread adoption.

40.2.1 Validity and Utility

The criteria for evaluating genetic tests are summarized by the four components of the ACCE analytic framework: analytical validity, clinical validity, clinical utility, and associated ethical, legal, and social implications [8]. Analytical validity refers to a test's ability to measure the genotype of interest accurately and reliably, which for germline pharmacogenetic variants is very robust [9,10]. More important is the appropriate selection of variants to interrogate for a particular drug response phenotype, also considering the ethnic group(s) that are being tested to maximize clinical validity.

Clinical validity refers to a test's ability to detect or predict the clinical disorder or phenotype associated with the genotype. Because most drug response phenotypes are multifactorial in origin, it is not always easy to achieve the high clinical validity for pharmacogenetic testing that is typically found in DNA testing for Mendelian disorders. Consequently, the positive predictive value of many pharmacogenetic assays can be low. For example, *CYP2C19*2* is a common variant allele (~15–25% allele frequency [11]) associated with high on-treatment platelet reactivity and increased risk for stent thrombosis, which is a rare clinical event (~0.5%) among clopidogrel-treated patients with acute coronary syndromes (ACSs) managed with percutaneous coronary intervention (PCI) [12–15]. These disparate allele and adverse clinical outcome frequencies result in a low positive predictive value for *CYP2C19* testing for these patients [16]. Many argue, however, that genetic testing in this scenario can still be useful and can help these patients avoid life-threatening and unnecessary risks, particularly when taken into consideration with other high-risk clinical factors. In this scenario, pharmacogenetic testing can be viewed as analogous to other nongenetic clinical variables with imperfect prediction (e.g., age, concurrent medications, comorbidities, liver function, etc.), but still provide useful and additive information [17,18].

The clinical utility of a test is a widely used measure of its usefulness in the clinic and in resulting changes in health outcomes. However, given the multidimensional nature of this kind of measurement, there is rarely consensus on precisely defining or adequately demonstrating it, particularly with regard to personalized medicine and pharmacogenetics [19]. The common benchmark for interventional evidence in medicine is a prospective randomized controlled trial, but such trials are often resource prohibitive for testing pharmacogenetic hypotheses, and they may be unethical to conduct for strong associations of severe adverse effects associated with high-risk genotypes. Nevertheless, two prospective randomized clinical trials have recently been completed that tested the utility of *CYP2C9* and *VKORC1* genotyping to guide warfarin dosing [20,21]. The results of these trials will likely greatly influence the future of pharmacogenetic testing for anticoagulation control and possibly other clinical scenarios with pharmacogenetic interactions. Still, in the absence of interventional clinical trial data, alternative evidence-gathering mechanisms are required, including incorporation of pharmacogenomics into premarket drug development, innovative clinical trial designs, and continued postmarket observational and mechanistic studies [22–24].

40.2.2 Education, Guidelines, and Recommendations

Clinician knowledge of pharmacogenetic testing significantly influences its successful integration into clinical practice. A previous study on medical genomics concluded that clinicians are generally not confident in providing genetic services because of insufficient training and knowledge [25]. Efforts are therefore needed to improve pharmacogenetic comprehension among clinicians, particularly by increasing its presence in medical school curriculums [26,27]. Moreover, the current significance of pharmacogenetics is not well understood by practicing clinicians, and they may not have adequate education on how to manage a patient based on pharmacogenetic test results. Consequently, there has been insufficient information and guidance available to help clinicians apply pharmacogenetic test results to individual patient management.

To address these educational needs, clinical practice guidelines that describe the utility of pharmacogenetic testing in different clinical scenarios are warranted to assist health professionals in evaluating and determining test utility

TABLE 40.1 Clinical Practice Guidelines, Recommendations, and Statements on Pharmacogenomic Testing

Organization	Drug(s)	Gene(s)	Reference
CPIC	Azathioprine/6-mercaptopurine	*TPMT*	Relling et al. [60,61]
	Clopidogrel	*CYP2C19*	Scott et al. [62,63]
	Warfarin	*CYP2C9/VKORC1*	Johnson et al. [64]
	Codeine	*CYP2D6*	Crews et al. [65]
	Abacavir	*HLA-B*	Martin et al. [66]
	Simvastatin	*SLCO1B1*	Wilke et al. [67]
	Allopurinal	*HLA-B*	Hershfield et al. [68]
	Tricyclic antidepressants (TCAs)	*CYP2D6/CYP2C19*	Hicks et al. [69]
	Carbamazepine	*HLA-B*	Leckband et al. [70]
	Fluoropyrimidine	*DPYD*	Caudle et al. [71]
ESF-UB	Abacavir	*HLA-B*5701*	Becquemont et al. [34]
	Azathioprine/6-mercaptopurine	*TPMT*	Becquemont et al. [34]
	Clopidogrel	*CYP2C19*	Becquemont et al. [34]
	Flucloxacillin	*HLA-B*5701*	Becquemont et al. [34]
	Statins	*SLCO1B1*	Becquemont et al. [34]
	Tacrolimus	*CYP3A5*	Becquemont et al. [34]
	Tamoxifen	*CYP2D6*	Becquemont et al. [34]
	Warfarin	*CYP2C9/VKORC1*	Becquemont et al. [34]
ACMG	Warfarin	*CYP2C9/VKORC1*	Flockhart et al. [72]
	Tamoxifen	*CYP2D6*	Lyon et al. [73]
EGAPP	Selective serotonin reuptake inhibitors (SSRIs)	*CYP2C19/CYP2D6*	EGAPP Working Group [31]
	Irinotecan	*UGT1A1*	EGAPP Working Group [32]
ACCF/AHA	Clopidogrel	*CYP2C19*	Holmes et al. [74]
KNMP-PWG	Multiple (53 drugs)	Multiple (11 genes)	Swen et al. [33]

ACCF—American College of Cardiology Foundation; ACMG—American College of Medical Genetics and Genomics; AHA—American Heart Association; CPIC—Clinical Pharmacogenetics Implementation Consortium; EGAPP—Evaluation of Genomic Applications in Practice and Prevention; ESF-UB—European Science Foundation-University of Barcelona (Pharmacogenetics and Pharmacogenomics: Practical Applications in Routine Medical Practice Conference); KNMP-PWG—Royal Dutch Association for the Advancement of Pharmacy-Pharmacogenetics Working Group.

and to facilitate the overall adoption of clinical pharmacogenetic testing [28]. Some of the first pharmacogenetic practice guidelines were developed by experts for *CYP2D6*- and *CYP2C19*-guided antidepressant dosing [29,30]; however, since then other organizations have published more formal documents (Table 40.1). Although some guidelines seek to evaluate evidence and establish whether testing is warranted or not [31,32], others do not make recommendations for or against testing and more so provide evidence-based recommended clinical actions when a patient's genotype is already known [33–35]. Notable examples include the thorough guidelines established by Royal Dutch Association for the Advancement of Pharmacy-Pharmacogenetics Working Group (KNMP-PWG), which reports on 53 drugs and 11 genes [33]; the expert-derived guidelines of the European Science Foundation (ESF) [34], and the evidence-based gene/drug guidelines of the Clinical Pharmacogenetics Implementation Consortium (CPIC) of the National Institutes of Health's Pharmacogenomics Research Network (PGRN), which are published and updated regularly on the Pharmacogenomics Knowledgebase (PharmGKB) website (www.pharmgkb.org) [35].

Evidence-based practice recommendation statements for pharmacogenetic testing have also been published by the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (Table 40.1) [36], launched by the Centers for Disease Control and Prevention (CDC) Office of Public Health Genomics. Although EGAPP previously found insufficient evidence to support recommendations for or against *CYP450* and *UGT1A1* genotyping in patients treated with selective serotonin reuptake inhibitors (SSRIs) and those treated with irinotecan, respectively, their reports did identify important research gaps that could facilitate future implementation [31,32]. Unfortunately, the level of evidence required for pharmacogenetic testing is not always clearly defined and is therefore often debated [18,37–40], as some view pharmacogenetics as an entity unique from other elements of genomic medicine which in many cases is already clinically warranted, whereas others unwearyingly await additional clinical evidence.

40.2.3 Pharmacogenetic Testing Regulation

Over the past several years, the FDA has revised numerous drug labels to now include relevant pharmacogenetic information; however, most do not require testing prior to initiating therapy. For example, CLIA-certified laboratory-developed pharmacogenetic tests and FDA-regulated tests without clinical claims do not necessarily have to provide full evidence of clinical validity and utility before being offered by a clinical laboratory. Although a number of important pharmacogenetic genes can be currently tested by CLIA-certified laboratories using laboratory-developed tests, there are only a small number of DNA-based pharmacogenetic tests that are actually FDA-approved for *in vitro* diagnostic testing at the present time, including assays for warfarin sensitivity (*CYP2C9* and *VKORC1*), *CYP2D6*, *CYP2C19*, and *UGT1A1*. Notably, the first sample-to-result system for genotyping *CYP2C19*2*, **3*, and **17* from a noninvasive cheek swab in under 60 minutes was very recently granted 510(k) clearance by the FDA.

For quality assurance, clinical laboratories also have the option of participating in the pharmacogenetic proficiency testing program sponsored by the College of American Pathologists (CAP), which provides graded and educational proficiency testing surveys [41]. Regarding test quality control and reference materials, the CDC Genetic Testing Reference Materials Coordination Program (GeT-RM), in collaboration with members of the pharmacogenetic testing community and Coriell Cell Repositories, have characterized a large panel of commercially available cell lines for genes and variants commonly included in pharmacogenetic testing panels and proficiency testing surveys [42]. An expanded, ongoing pharmacogenomics reference material project by the GeT-RM is currently interrogating a larger panel of genes and variants using single gene-targeted assays, genotyping microarrays, and next-generation sequencing of custom-designed gene panels.

40.2.4 Pharmacogenetic Testing Turnaround Time

Although rapid genetic testing can be performed in some clinical scenarios, typical turnaround time for genetic testing in a clinical laboratory is usually days to weeks, depending on the testing technology. Unfortunately, for the majority of currently actionable pharmacogenetic variants, this is unacceptable for efficient implementation into routine clinical care. For example, in the cardiovascular pharmacogenetics field, *CYP2C9*/*VKORC1* and *CYP2C19* genotype data are most effective at the time of drug initiation for both warfarin and clopidogrel, respectively [43]. Moreover, given the demanding environments common to most anticoagulation clinics and cardiac catheterization laboratories, disruption of routine care by interfacing with an external clinical laboratory that has additional genetic testing logistics and unique laboratory information management systems can present further complexities for effective use of pharmacogenetic testing.

40.3 POINT-OF-CARE PHARMACOGENETIC TESTING

To address the issue of testing turnaround time for more efficient implementation, many commercial companies have been developing genotyping platforms that offer rapid sample-to-result assays that will be highly beneficial for integrating pharmacogenetics at the point of care [9,44,45]. A notable example comes from the Reassessment of Anti-Platelet Therapy Using An Individualized Strategy Based on Genetic Evaluation (RAPID GENE) trial, which recently reported use of *CYP2C19*2* (c.681G>A) point-of-care genetic testing for cardiac patients initiating clopidogrel therapy following PCI [46]. In this proof-of-concept study, patients were randomly assigned to rapid point-of-care genotyping or to standard treatment; those in the genotyping group were tested for *CYP2C19*2* using a cheek swab genetic testing device that reported results within 60 minutes. Carriers were treated with prasugrel, and noncarriers and patients in the standard treatment group were treated with clopidogrel. Importantly, no carriers in the rapid genotyping group had high on-treatment platelet reactivity (HTPR) at day seven (the primary end point), compared with 30% of patients given standard treatment ($p = 0.0092$).

Although this study showed that point-of-care genetic testing following PCI can be performed effectively by nursing staff, and that personalized antiplatelet therapy can reduce HTPR in this patient population, it has yet to be established that this testing strategy results in better clinical outcomes. To address this important question, a number of related trials are currently ongoing that measure clinical outcomes following rapid and/or point-of-care *CYP2C19*2* testing. For example, RAPID STEMI (Reassessment of Anti-Platelet Therapy Using an Individualized Strategy in Patients with ST-segment Elevation Myocardial Infarction; ClinicalTrials.gov Identifier: NCT01452139) is evaluating the feasibility, efficacy, and safety of pharmacogenetic-guided antiplatelet therapy for the treatment of STEMI patients following PCI using point-of-care *CYP2C19*2*, **17*, and *ABCB1* 3435C>T genetic testing. Although the primary outcome is high on-treatment platelet reactivity, there are a number of secondary outcomes of clinical measures (e.g., composite of death from cardiovascular causes, nonfatal myocardial infarction, and re-hospitalization; bleeding risk).

TAILOR-PCI (Tailored Antiplatelet Therapy Following PCI; ClinicalTrials.gov Identifier: NCT01742117) is a large trial with an estimated enrollment of almost 6000 patients that is testing the hypothesis that, following PCI and using a *CYP2C19* genotyping strategy, ticagrelor is superior to clopidogrel in reducing a composite end point of major adverse cardiovascular events (MACE). The primary outcome is the occurrence of MACE within one year following PCI, and recruitment is scheduled until 2016. PRU-MATRIX (Customized Choice of P2Y12 Oral Receptor Blocker Based on Phenotype Assessment via Point of Care Testing; ClinicalTrials.gov Identifier: NCT01477775) is another ongoing clinical trial that is comparing standard of care to a customized antiplatelet algorithm that includes both platelet reactivity and genetic (*CYP2C19*) variables. The primary outcome is the time to occurrence of cardiovascular death, myocardial infarction, stroke or bleeding, and the proportion of patients in therapeutic range for platelet aggregation (P2Y12 pathway activity according to PRU values).

Despite the enthusiasm for rapid-turnaround time pharmacogenetic testing and the success of the RAPID GENE trial, issues remain when one considers transferring this type of genetic testing from research study to routine clinical care. For example, the regulatory landscape of point-of-care testing, particularly that involving genetic testing, can be complicated. The recent FDA approval of the device used in the RAPID GENE trial does not support its use specifically as "point of care" because of the need for personnel with adequate genetics training. Instead, it is approved and marketed as a "rapid sample-to-result" platform.

Point-of-care testing is, by definition, clinical laboratory testing performed at or near the site of clinical care delivery by personnel (or patients) whose primary training is not in the clinical laboratory sciences. The pathways for FDA approval of point-of-care devices include either 510(k) clearance, premarket approval applications, or CLIA waivers when a device has a negligible likelihood of erroneous results and has no risk of harm if performed incorrectly. Point-of-care pharmacogenetic testing is likely not amenable to a CLIA waiver, which highlights a potential challenge when performing clinical genetic testing at the point of care in the absence of personnel with certified genetics expertise. Depending on regional regulations, it is possible that a local CLIA-certified genetics laboratory may be required to oversee point-of-care testing by remotely managing interpretation, performance, quality control/assurance, and participation in relevant proficiency testing programs. This could increase the complexity and overhead costs of point-of-care testing and add potential difficulties when defining the relationship between point-of-care clinical staff and CLIA-certified genetic laboratories.

A technical challenge for point-of-care pharmacogenetic testing involves the content of the genotyping assays themselves. The RAPID GENE trial was centered on a single polymorphic allele, which for future pharmacogenetic assays will not be adequate. This is even more relevant when deploying point-of-care pharmacogenetic testing across more diverse patient populations, as the frequencies of relevant variant alleles differ between racial groups and ethnicities. For example, the *CYP2C19*4B* loss-of-function allele has a frequency of ~2% among Ashkenazi Jewish individuals, but the frequency is lower in other racial and ethnic groups [11,47]. In addition, more genes and functional variants are necessary for some currently actionable pharmacogenetic examples (e.g., warfarin); also, ongoing genome sequencing studies are likely to identify more variants with appreciable effect sizes that will justify inclusion in future point-of-care testing panels.

Some of the more robust pharmacogenetic associations at the present time involve specific human leukocyte antigen (HLA) alleles of the major histocompatibility complex (e.g., *HLA-B*5701* and flucloxacillin-induced liver injury and abacavir-induced hypersensitivity [48]; *HLA-B*1502*/*HLA-A*3101* and carbamazepine-induced hypersensitivity [49,50]). Unfortunately, *HLA* genotyping is one of the more challenging molecular assays, requiring combinatorial multiplexing that is beyond the technical capacity of current point-of-care platforms. Ideally, future point-of-care genetic testing platforms will overcome the technical need for multiplexed sample-to-answer genotyping and be able to include larger variant panels with content specifically selected for clinically actionable alleles.

40.4 PRE-EMPTIVE PHARMACOGENETIC TESTING

Another recently actualized strategy for clinical pharmacogenetics that can circumvent some of the issues surrounding both traditional external laboratory and point-of-care testing is pre-emptive clinical genotyping. Although this model has its own set of obstacles for effective clinical implementation, pre-emptive pharmacogenetic testing programs are currently underway at selected academic medical centers [51–54]. This approach deposits clinical genotype data in patient records pre-emptively, usually through coordinated prospective sampling and/or biobanking, CLIA-certified testing and informatics, and alerts prescribers at the point of care through sophisticated electronic clinical decision support (CDS) systems when a drug is ordered for a patient with an at-risk genotype. The immediate knowledge of personalized and relevant pharmacogenetic variation, with interpretation and possible/recommended actions without disruption of routine clinical care, is the clear advantage to this strategy.

Additionally, pre-emptive genotyping can still utilize CLIA-certified laboratory testing, but without the usual concerns regarding turnaround time because it is done prior to drug prescription and possibly before clinical indication. However, the necessary institutional investments in informatics, CDS, healthcare provider participation and education, and infrastructure for sampling/biobanking and testing in a CLIA-certified environment suggest that this exciting mode of clinical pharmacogenetic delivery will be limited to large academic medical centers. Although not meant as a comprehensive listing, an overview of selected pre-emptive pharmacogenetic testing programs follows.

40.4.1 Vanderbilt University Medical Center: PREDICT

One example of pre-emptive pharmacogenetic testing is the Pharmacogenomic Resource for Enhanced Decisions in Care and Treatment (PREDICT) program at the Vanderbilt University Medical Center (VUMC) [51]. Launched in 2010, the initial phase of this program involves ~3000 patients scheduled for cardiac catheterization, coupled with pre-emptive genotyping with a multiplexed platform that includes *CYP2C19*. Genotype data are deposited into the electronic medical record (EMR), and point-of-care CDS is deployed when clopidogrel is prescribed for patients with variant *CYP2C19* genotypes (intermediate and poor metabolizers). The evidence for *CYP2C19* genotype and clopidogrel response is generally considered strong and amenable to initial implementation; however, the long-term goal of PREDICT is to establish a framework and comprehensive infrastructure for pre-emptive incorporation into the VUMC EMR of genomic and other high-dimensional patient-specific data. Notably, this program has highlighted many of the important issues surrounding implementation of pharmacogenetic testing, including the importance of engaging patients and relevant physician providers, the necessary institutional commitments, the multidisciplinary effort required in reviewing evidence and developing CDS, and the need for genotyping in a CLIA-certified laboratory. PREDICT is currently being expanded to include warfarin/*CYP2C9*-*VKORC1* and simvastatin/*SLCO1B1* drug/gene pairs.

40.4.2 University of Florida and Shands Hospital: PMP

The Personalized Medicine Program (PMP), of the University of Florida (UF) and Shands Hospital was launched in 2011 as a translational clinical implementation program within the UF Clinical and Translational Science Institute [54]. Its goal is to address some of the challenges of genomic medicine, including how to store and utilize genomic data for clinical use, define genetic data that are sufficiently robust for clinical use, and develop CDS to guide clinicians in the use of genomic data. A multidisciplinary PMP subcommittee evaluates the pharmacogenetics literature to identify clinical examples where there is sufficiently robust evidence to warrant clinical implementation. Like PREDICT, the initial PMP pilot, deployed in 2012, is also centered on clopidogrel and *CYP2C19*. However, its pre-emptive multiplexed genotyping panel includes a much larger number of genes and variants [55], tested under research consent for storage in the EHR and for possible use in future implementation efforts. This custom-designed panel, tested in a CLIA-certified laboratory, includes 252 pharmacogenetic variants and was designed in collaboration with PharmGKB (www.pharmgkb.org). Future plans for the UF PMP include expanding the clopidogrel/*CYP2C19* program to other private practices across Florida and implementing additional pharmacogenetic examples with CPIC guidelines in other practice settings (e.g., pediatric hematology/oncology).

40.4.3 University of Chicago: 1200 Patients Project

The 1200 Patients Project, recently deployed by the University of Chicago, is another pre-emptive pharmacogenetic genotyping program [52]. Here, patients receiving outpatient medical care who are taking one to six prespecified prescription medications under the care of recruited physicians (two cardiologists, one pulmonologist, one hepatologist, one gastroenterologist, three oncologists, and four general internists) are prospectively genotyped in a CLIA-certified laboratory with a large panel of clinically relevant germline pharmacogenetic variants. The panel provides comprehensive genotype results for a large number of variants known to influence response and/or toxicity of commonly used drugs. Patient-specific results are subsequently made available exclusively to the enrolling provider through an online genomic prescribing system [56]. Like other pre-emptive programs, the initial phase of the project is centered on the return of pharmacogenetic information and the overall examination of its use. Whether physicians take this information into consideration during clinic visits, and whether it results in altered patterns of prescribing in patients who are genetically at high risk of adverse drug outcomes/nonresponse, are the primary and secondary end points.

40.4.4 St. Jude Children's Research Hospital: PG4KDS

The St. Jude PG4KDS pharmacogenetics implementation program is migrating genotype data derived from a large microarray-based panel (1936 genomic variants in 225 genes) [57] that have been tested in a CLIA-certified laboratory, into the EMR as prospective patients are enrolled [58]. The patient population of the program are enrolled through pediatric clinics and are typically receiving multiple therapies. Like other programs, the majority of the genotype data are maintained in a research environment, with only selected genes and variants classified as priorities when there is sufficient evidence for deposit into the EMR. The EMR subsequently deploys CDS when an intersection between a prescription and an at-risk genotype occurs. Current priority genes being implemented by PG4KDS include *CYP2C19*, *CYP2D6*, *SLCO1B1*, and *TPMT*. These genotype results are aggregated into a single section in the EMR on a per-patient basis and can be easily expanded to accommodate additional genes and drugs as they are identified.

40.4.5 Icahn School of Medicine at Mount Sinai: CLIPMERGE–PGx

The CLIPMERGE–PGx (Clinical Implementation of Personalized Medicine through Electronic Health Records and Genomics—Pharmacogenomics) program of the Icahn School of Medicine in New York, is a pre-emptive pharmacogenetics initiative that is genotyping a panel of germline variants in reconsented patients initially enrolled in the school's Bio*Me* Biobank Program [53]. Notably, the CLIPMERGE platform is an independent and external system that communicates with the EHR and houses decision support rules based on actionable variants and relevant phenotypic data from the CLIPMERGE database, including longitudinal clinical data extracted from the EHR. A Clinical Risk Assessment Engine (CRAE) is included in this platform with a rules engine that relates actionable genotype–phenotype pairs to genome-informed advice messages. If predefined rules are met, decision support based on available CPIC guidelines is delivered in real-time at the point of care through the EHR. Patients for the pilot program are selected based on their regular attendance at the Mount Sinai Internal Medicine Associates (IMA) clinic for their primary care. In order to enrich the program with patients likely to have pharmacogenetic interactions, targeted patients include those currently taking clopidogrel, warfarin, simvastatin, tricyclic antidepressants (TCAs), and/or selective serotonin reuptake inhibitors (SSRIs). Consented patients provide a saliva sample for genetic testing in a CLIA-certified laboratory and genotype data are ultimately stored in the external CLIPMERGE database. Consistent with other programs, transactional data regarding CDS metrics, usage, and clinical parameters are collected on an ongoing basis from the EHR and CLIPMERGE CDS system. The CLIPMERGE PGx pilot program has 1500 consenting Bio*Me* participants; however, the eventual aim is to recruit all participants from Bio*Me* for further pharmacogenetic and other genomic medicine implementation initiatives.

40.4.6 PGRN: Translational Pharmacogenetics Program

In an effort to better identify and possibly address some of the barriers to implementing pharmacogenetics in a collaborative multi-institution effort, the PGRN (Pharmacogenomics Research Network) recently established the Translational Pharmacogenomics Program, which comprises six diverse healthcare systems (some detailed earlier): the University of Maryland, the University of Florida, St. Jude Children's Research Hospital, Vanderbilt University, the Mayo Clinic, and Ohio State University [59]. PharmGKB (www.pharmgkb.org) serves in a coordination and knowledge sharing/dissemination role. For this initiative, each site is implementing one or more pharmacogenetic tests, adapting evidence-based CPIC guidelines to local policies and practices, and identifying and attempting to overcome real-life policy, logistic, and translational barriers. Notably, both point-of-care testing (e.g., *CYP2C19* testing in PCI patients) and pre-emptive models are being executed and studied. Each implementation step will be systematically evaluated to develop a practical evidence-based toolbox of best practices for pharmacogenetics implementation that will be useful across a wide spectrum of healthcare systems. These tools will be disseminated through publication in peer-reviewed journals and updated regularly on the PharmGKB website. Other instruments that will be developed for dissemination include a best practices manual for implementation of pharmacogenetics testing, and questionnaires and surveys to assess implementation metrics and effectiveness.

40.5 CONCLUSION AND FUTURE PERSPECTIVES

The ongoing advances in point-of-care and pre-emptive genetic testing are greatly facilitating the implementation of efficient and effective clinical pharmacogenetics. Both strategies circumvent the barrier of test turnaround time and, therefore, provide the opportunity to integrate pharmacogenetic test results at the point of care. Although these strategies are very promising, the increasing accessibility of whole-genome sequencing suggests that the long-term future landscape may no longer have a need for targeted genotyping if patients already have their sequence data available. However, for the shorter-term future, the use of point-of-care genetic testing will undoubtedly be very useful in the clinic (if accepted by regulatory agencies and laboratory providers) and for prospective genetic intervention trials.

Pre-emptive pharmacogenetic testing is an exciting strategy for delivering genetic-based CDS at the point of care to help guide pharmacotherapy; however, the generalizability of this strategy beyond large academic medical centers may be limited by the necessary institutional investments, information technology infrastructure, and broad provider support. Regardless of how the pharmacogenetic data are derived (or when), of utmost importance will be to continue the necessary efforts toward provider and patient education, accessible and appropriate CDS, and rigorous assessment of clinical utility by the pharmacogenetics community.

REFERENCES

[1] Kan YW, Dozy AM. Antenatal diagnosis of sickle-cell anaemia by D.N.A. analysis of amniotic-fluid cells. Lancet 1978;2(8096):910–2.

[2] Scott SA, Edelmann L, Liu L, Luo M, Desnick RJ, Kornreich R. Experience with carrier screening and prenatal diagnosis for 16 Ashkenazi Jewish genetic diseases. Hum Mutat 2010;31(11):1240–50.

[3] Ng PC, Murray SS, Levy S, Venter JC. An agenda for personalized medicine. Nature 2009;461(7265):724–6.

[4] Motulsky AG. Drug reactions enzymes, and biochemical genetics. J Am Med Assoc 1957;165(7):835–7.

[5] Scott SA. Personalizing medicine with clinical pharmacogenetics. Genet Med 2011;13(12):987–95.

[6] Scott SA. Clinical pharmacogenomics: opportunities and challenges at point of care. Clin Pharmacol Ther 2013;93(1):33–5.

[7] Ryan SG. Regression to the truth: replication of association in pharmacogenetic studies. Pharmacogenomics 2003;4(2):201–7.

[8] Sanderson S, Zimmern R, Kroese M, Higgins J, Patch C, Emery J. How can the evaluation of genetic tests be enhanced? Lessons learned from the ACCE framework and evaluating genetic tests in the United Kingdom. Genet Med 2005;7(7):495–500.

[9] Koch WH. Technology platforms for pharmacogenomic diagnostic assays. Nat Rev Drug Discov 2004;3(9):749–61.

[10] King CR, Porche-Sorbet RM, Gage BF, Ridker PM, Renaud Y, Phillips MS, et al. Performance of commercial platforms for rapid genotyping of polymorphisms affecting warfarin dose. Am J Clin Pathol 2008;129(6):876–83.

[11] Martis S, Peter I, Hulot JS, Kornreich R, Desnick RJ, Scott SA. Multi-ethnic distribution of clinically relevant CYP2C genotypes and haplotypes. Pharmacogenomics J 2013;13(4):369–77.

[12] Brandt JT, Close SL, Iturria SJ, Payne CD, Farid NA, Ernest 2nd CS, et al. Common polymorphisms of CYP2C19 and CYP2C9 affect the pharmacokinetic and pharmacodynamic response to clopidogrel but not prasugrel. J Thromb Haemost 2007;5(12):2429–36.

[13] Shuldiner AR, O'Connell JR, Bliden KP, Gandhi A, Ryan K, Horenstein RB, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. JAMA 2009;302(8):849–57.

[14] Hulot JS, Collet JP, Silvain J, Pena A, Bellemain-Appaix A, Barthélémy O, et al. Cardiovascular risk in clopidogrel-treated patients according to cytochrome P450 2C19*2 loss-of-function allele or proton pump inhibitor coadministration: a systematic meta-analysis. J Am Coll Cardiol 2010;56(2):134–43.

[15] Mega JL, Simon T, Collet JP, Anderson JL, Antman EM, Bliden K, et al. Reduced-function CYP2C19 genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI: a meta-analysis. JAMA 2010;304(16):1821–30.

[16] Hulot JS, Fuster V. Antiplatelet therapy: personalized medicine for clopidogrel resistance? Nat Rev Cardiol 2009;6(5):334–6.

[17] Green MJ, Botkin JR. "Genetic exceptionalism" in medicine: clarifying the differences between genetic and nongenetic tests. Ann Intern Med 2003;138(7):571–5.

[18] Relling MV, Altman RB, Goetz MP, Evans WE. Clinical implementation of pharmacogenomics: overcoming genetic exceptionalism. Lancet Oncol 2010;11(6):507–9.

[19] Lesko LJ, Zineh I, Huang SM. What is clinical utility and why should we care? Clin Pharmacol Ther 2010;88(6):729–33.

[20] Kimmel SE, French B, Anderson JL, Gage BF, Johnson JA, Rosenberg YD, et al. Rationale and design of the Clarification of Optimal Anticoagulation through Genetics trial. Am Heart J 2013;166(3):435–41. e2

[21] van Schie RM, Wadelius MI, Kamali F, Daly AK, Manolopoulos VG, de Boer A, et al. Genotype-guided dosing of coumarin derivatives: the European pharmacogenetics of anticoagulant therapy (EU-PACT) trial design. Pharmacogenomics 2009;10(10):1687–95.

[22] Zineh I, Lesko L. Pharmacogenetics in medicine: barriers, critical factors and a framework for dialogue. Pers Med 2009;6(4):359–61.

[23] Burns DK. Developing pharmacogenetic evidence throughout clinical development. Clin Pharmacol Ther 2010;88(6):867–70.

[24] Woodcock J. Assessing the clinical utility of diagnostics used in drug therapy. Clin Pharmacol Ther 2010;88(6):765–73.

[25] Scheuner MT, Sieverding P, Shekelle PG. Delivery of genomic medicine for common chronic adult diseases: a systematic review. JAMA 2008;299(11):1320–34.

[26] Gurwitz D, Lunshof JE, Dedoussis G, Flordellis CS, Fuhr U, Kirchheiner J, et al. Pharmacogenomics education: International Society of Pharmacogenomics recommendations for medical, pharmaceutical, and health schools deans of education. Pharmacogenomics J 2005;5(4):221–5.

[27] Green JS, O'Brien TJ, Chiappinelli VA, Harralson AF. Pharmacogenomics instruction in US and Canadian medical schools: implications for personalized medicine. Pharmacogenomics 2010;11(9):1331–40.

[28] Amstutz U, Carleton BC. Pharmacogenetic testing: time for clinical practice guidelines. Clin Pharmacol Ther 2011;89(6):924–7.

[29] Kirchheiner J, Brøsen K, Dahl ML, Gram LF, Kasper S, Roots I, et al. CYP2D6 and CYP2C19 genotype-based dose recommendations for antidepressants: a first step towards subpopulation-specific dosages. Acta Psychiatr Scand 2001;104(3):173–92.

[30] de Leon J, Armstrong SC, Cozza KL. Clinical guidelines for psychiatrists for the use of pharmacogenetic testing for CYP450 2D6 and CYP450 2C19. Psychosomatics 2006;47(1):75–85.

[31] EGAPP Working Group. Recommendations from the EGAPP Working Group: testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. Genet Med 2007;9(12):819–25.

[32] Wrishko RE, Ernest 2nd CS, Small DS, Li YG, Weerakkody GJ, Riesmeyer JR, et al. Population pharmacokinetic analyses to

evaluate the influence of intrinsic and extrinsic factors on exposure of prasugrel active metabolite in TRITON-TIMI 38. J Clin Pharmacol 2009;49(8):984–98.

[33] Swen JJ, Nijenhuis M, de Boer A, Grandia L, Maitland-van der Zee AH, Mulder H, et al. Pharmacogenetics: from bench to byte – an update of guidelines. Clin Pharmacol Ther 2011;89(5):662–73.

[34] Becquemont L, Alfirevic A, Amstutz U, Brauch H, Jacqz-Aigrain E, Laurent-Puig P, et al. Practical recommendations for pharmacogenomics-based prescription: 2010 ESF-UB Conference on Pharmacogenetics and Pharmacogenomics. Pharmacogenomics 2011;12(1):113–24.

[35] Relling MV, Klein TE. CPIC: clinical pharmacogenetics implementation consortium of the pharmacogenomics research network. Clin Pharmacol Ther 2011;89(3):464–7.

[36] Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The evaluation of genomic applications in practice and prevention (EGAPP) initiative: methods of the EGAPP working group. Genet Med 2009;11(1):3–14.

[37] Grosse SD, McBride CM, Evans JP, Khoury MJ. Personal utility and genomic information: look before you leap. Genet Med 2009;11(8):575–6.

[38] Khoury MJ. Dealing with the evidence dilemma in genomics and personalized medicine. Clin Pharmacol Ther 2010;87(6):635–8.

[39] Altman RB. Pharmacogenomics: “noninferiority” is sufficient for initial implementation. Clin Pharmacol Ther 2011;89(3):348–50.

[40] Khoury MJ, Gwinn M, Dotson WD, Bowen MS. Is there a need for PGxceptionalism? Genet Med 2011;13(10):866–7.

[41] Wu AH. Genotype and phenotype concordance for pharmacogenetic tests through proficiency survey testing. Arch Pathol Lab Med 2013;137(9):1232–6.

[42] Pratt VM, Zehnbauer B, Wilson JA, Baak R, Babic N, Bettinotti M, et al. Characterization of 107 genomic DNA reference materials for CYP2D6, CYP2C19, CYP2C9, VKORC1, and UGT1A1: a GeT-RM and association for molecular pathology collaborative project. J Mol Diagn 2010;12(6):835–46.

[43] Collet JP, Hulot JS, Anzaha G, Pena A, Chastre T, Caron C, et al. High doses of clopidogrel to overcome genetic resistance: the randomized crossover CLOVIS-2 (clopidogrel and response variability investigation study 2). JACC Cardiovasc Interv 2011;4(4):392–402.

[44] Dobson MG, Galvin P, Barton DE. Emerging technologies for point-of-care genetic testing. Expert Rev Mol Diagn 2007;7(4):359–70.

[45] Buchan BW, Peterson JF, Cogbill CH, Anderson DK, Ledford JS, White MN, et al. Evaluation of a microarray-based genotyping assay for the rapid detection of cytochrome P450 2C19 *2 and *3 polymorphisms from whole blood using nanoparticle probes. Am J Clin Pathol 2011;136(4):604–8.

[46] Roberts JD, Wells GA, Le May MR, Labinaz M, Glover C, Froeschl M, et al. Point-of-care genetic testing for personalisation of antiplatelet treatment (RAPID GENE): a prospective, randomised, proof-of-concept trial. Lancet 2012;379(9827):1705–11.

[47] Scott SA, Martis S, Peter I, Kasai Y, Kornreich R, Desnick RJ. Identification of CYP2C19*4B: pharmacogenetic implications for drug metabolism including clopidogrel responsiveness. Pharmacogenomics J 2012;12(4):297–305.

[48] Mallal S, Phillips E, Carosi G, Molina JM, Workman C, Tomazic J, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008;358(6):568–79.

[49] Chen P, Lin JJ, Lu CS, Ong CT, Hsieh PF, Yang CC, et al. Carbamazepine-induced toxic effects and HLA-B*1502 screening in Taiwan. N Engl J Med 2011;364(12):1126–33.

[50] McCormack M, Alfirevic A, Bourgeois S, Farrell JJ, Kasperavičiūtė D, Carrington M, et al. HLA-A*3101 and carbamazepine-induced hypersensitivity reactions in Europeans. N Engl J Med 2011;364(12):1134–43.

[51] Pulley JM, Denny JC, Peterson JF, Bernard GR, Vnencak-Jones CL, Ramirez AH, et al. Operational implementation of prospective genotyping for personalized medicine: the design of the Vanderbilt PREDICT project. Clin Pharmacol Ther 2012;92(1):87–95.

[52] O’Donnell PH, Bush A, Spitz J, Danahey K, Saner D, Das S, et al. The 1200 patients project: creating a new medical model system for clinical implementation of pharmacogenomics. Clin Pharmacol Ther 2012;92(4):446–9.

[53] Gottesman O, Scott SA, Ellis SB, Overby CL, Ludtke A, Hulot JS, et al. The CLIPMERGE PGx Program: clinical implementation of personalized medicine through electronic health records and genomics-pharmacogenomics. Clin Pharmacol Ther 2013;94(2):214–7.

[54] Johnson JA, Elsey AR, Clare-Salzler MJ, Nessl D, Conlon M, Nelson DR. Institutional profile: University of Florida and Shands Hospital Personalized Medicine Program: clinical implementation of pharmacogenetics. Pharmacogenomics 2013;14(7):723–6.

[55] Johnson JA, Burkley BM, Langaee TY, Clare-Salzler MJ, Klein TE, Altman RB. Implementing personalized medicine: development of a cost-effective customized pharmacogenetics genotyping array. Clin Pharmacol Ther 2012;92(4):437–9.

[56] Ratain MJ. Personalized medicine: building the GPS to take us there. Clin Pharmacol Ther 2007;81(3):321–2.

[57] Fernandez CA, Smith C, Yang W, Lorier R, Crews KR, Kornegay N, et al. Concordance of DMET plus genotyping results with those of orthogonal genotyping methods. Clin Pharmacol Ther 2012;92(3):360–5.

[58] Hicks JK, Crews KR, Hoffman JM, Kornegay NM, Wilkinson MR, Lorier R, et al. A clinician-driven automated system for integration of pharmacogenetic interpretations into an electronic medical record. Clin Pharmacol Ther 2012;92(5):563–6.

[59] Shuldiner AR, Relling MV, Peterson JF, Hicks JK, Freimuth RR, Sadee W, et al. The Pharmacogenomics Research Network Translational Pharmacogenetics Program: overcoming challenges of real-world implementation. Clin Pharmacol Ther 2013;94(2):207–10.

[60] Relling MV, Gardner EE, Sandborn WJ, Schmiegelow K, Pui CH, Yee SW, et al. Clinical pharmacogenetics implementation consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing. Clin Pharmacol Ther 2011;89(3):387–91.

[61] Relling MV, Gardner EE, Sandborn WJ, Schmiegelow K, Pui CH, Yee SW, et al. Clinical pharmacogenetics implementation consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing: 2013 update. Clin Pharmacol Ther 2013;93(4):324–5.

[62] Scott SA, Sangkuhl K, Gardner EE, Stein CM, Hulot JS, Johnson JA, et al. Clinical pharmacogenetics implementation consortium guidelines for cytochrome P450–2C19 (CYP2C19) genotype and clopidogrel therapy. Clin Pharmacol Ther 2011;90(2):328–32.

[63] Scott SA, Sangkuhl K, Stein CM, Hulot JS, Mega JL, Roden DM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. Clin Pharmacol Ther 2013;94(3):317–23.

[64] Johnson JA, Gong L, Whirl-Carrillo M, Gage BF, Scott SA, Stein CM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. Clin Pharmacol Ther 2011;90(4):625–9.

[65] Crews KR, Gaedigk A, Dunnenberger HM, Klein TE, Shen DD, Callaghan JT, et al. Clinical pharmacogenetics implementation

consortium (CPIC) guidelines for codeine therapy in the context of cytochrome P450 2D6 (CYP2D6) genotype. Clin Pharmacol Ther 2012;91(2):321–6.

[66] Martin MA, Klein TE, Dong BJ, Pirmohamed M, Haas DW, Kroetz DL, et al. Clinical pharmacogenetics implementation consortium guidelines for HLA-B genotype and abacavir dosing. Clin Pharmacol Ther 2012;91(4):734–8.

[67] Wilke RA, Ramsey LB, Johnson SG, Maxwell WD, McLeod HL, Voora D, et al. The clinical pharmacogenomics implementation consortium: CPIC guideline for SLCO1B1 and simvastatin-induced myopathy. Clin Pharmacol Ther 2012;92(1):112–7.

[68] Hershfield MS, Callaghan JT, Tassaneeyakul W, Mushiroda T, Thorn CF, Klein TE, et al. Clinical pharmacogenetics implementation consortium guidelines for human leukocyte antigen-B genotype and allopurinol dosing. Clin Pharmacol Ther 2013;93(2):153–8.

[69] Hicks JK, Swen JJ, Thorn CF, Sangkuhl K, Kharasch ED, Ellingrod VL, et al. Clinical pharmacogenetics implementation consortium guideline for CYP2D6 and CYP2C19 genotypes and dosing of tricyclic antidepressants. Clin Pharmacol Ther 2013;93(5):402–8.

[70] Leckband SG, Kelsoe JR, Dunnenberger HM, George Jr AL, Tran E, Berger R, et al. Clinical pharmacogenetics implementation consortium guidelines for HLA-B genotype and carbamazepine dosing. Clin Pharmacol Ther 2013;94(3):324–8.

[71] Caudle KE, Thorn CF, Klein TE, Swen JJ, McLeod HL, Diasio RB, et al. Clinical pharmacogenetics implementation consortium guidelines for dihydropyrimidine dehydrogenase genotype and fluoropyrimidine dosing. Clin Pharmacol Ther 2013;94(6):640–5.

[72] Flockhart DA, O'Kane D, Williams MS, Watson MS, Flockhart DA, Gage B, et al. Pharmacogenetic testing of CYP2C9 and VKORC1 alleles for warfarin. Genet Med 2008;10(2):139–50.

[73] Lyon E, Gastier Foster J, Palomaki GE, Pratt VM, Reynolds K, Sábato MF, et al. Laboratory testing of CYP2D6 alleles in relation to tamoxifen therapy. Genet Med 2012;14(12):990–1000.

[74] Holmes Jr DR, Dehmer GJ, Kaul S, Leifer D, O'Gara PT, Stein CM. ACCF/AHA clopidogrel clinical alert: approaches to the FDA "boxed warning": a report of the American College of Cardiology Foundation Task Force on clinical expert consensus documents and the American Heart Association endorsed by the Society for Cardiovascular Angiography and Interventions and the Society of Thoracic Surgeons. J Am Coll Cardiol 2010;56(4):321–41.

Chapter 41

Ethical Considerations in Pharmacogenomic Testing and Research in Pediatrics

Chi Dang Hornik,[a] Cait O'Donnell,[b] and Raymond Barfield[c]

[a]*Neonatal Intensive Care, Duke Children's Hospital & Health Center, Duke University Medical Center, USA,* [b]*Ethics in Military Medicine Research Group, McMaster University, UK,* [c]*Pediatric Quality of Life Program, Duke University, USA*

Chapter Outline

41.1 Overview 931
41.2 Historical Perspective 933
41.3 Ethical Significance of Benefit, Risk, and Uncertainty in Pediatric Pharmacogenomics 935
41.3.1 What is "Consent" in Pediatrics? 938
41.3.2 Assent in Pediatric Research 940
41.3.3 Minimal Risk 943
41.4 Federal Regulations and National Healthcare Priorities Related to Pharmacogenomics 943
41.5 Translation of Pharmacogenomics into the Clinical Setting 945
41.6 Future Implications of Pharmacogenomics Testing in Children 946
41.7 Conclusion 948
References 948

41.1 OVERVIEW

Enterprises such as the Human Genome Project in 2003 and the International HapMap Project in 2007 have advanced our understanding of important genetic associations to drug responses. Growing from this new understanding, pharmacogenomic (PGx) testing is increasingly studied and used in children to improve pediatric dosing strategies. Pharmacogenomics is the subcategory of genomics research that has perhaps the most immediate potential impact in pediatric clinical trials and therapy. This area has enormous potential to benefit children, but also creates possibilities for social and economic discrimination, as well as abuse of information in genomic databases.

Pharmacogenomics combines human genomics with pharmacogenetics to study the role of the individual's genetic makeup in his or her response to a drug [1]. The promise of this approach is that patients can receive individually tailored medical therapy and dosing. This technology advances the important concept of risk stratification, which has historically led to better protection of children with lower-risk diseases from excess toxicity, at the same time ensuring that children with higher-risk diseases are not undertreated.

In identifying the specific genetic determinants involved in drug disposition, metabolism, and response throughout a child's development, the goal is for pediatricians and clinical investigators to be better able to choose the best drug for that child and to more accurately determine safe and effective dosing. However, as PGx technology evolves in pediatrics, several key ethical issues unique to children will need to be considered. Both ethical and political issues related to PGx testing have been extensively addressed in the adult population, and the issues that are raised in this chapter are relevant to adults for many reasons, including the difficulty of understanding the true implications of the testing. But when the patient or research participant is a child, these issues become more complex, in part because pediatric participants range from neonates who do not participate in decision making to older adolescents who often participate more but without the complete autonomy and legal status of adults. It should be noted that some of the ethical and regulatory issues presented in this chapter are specific to the American health system. It should also be

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00041-4
© 2014 Elsevier Inc. All rights reserved.

emphasized that these challenges are not static but change as a child grows, as our knowledge of PGx evolves, and as legislative responses mature in response to these new technologies.

In this chapter, we explore the historical background against which clinical and research ethics have developed as well some of the more important approaches to framing ethical questions in pediatric clinical practice and research. We then focus on three broad areas of ethical concern that arise in pediatrics. First, should PGx testing be performed in children and, if yes, what are the specific benefits and risks to the child? Second, what are the ethical considerations of pediatric PGx testing in the context of research and how are they similar to those in other forms of pediatric research in terms of informed consent, assent, and research risk assessment? Third, how do these general ethical concerns provide a foundation and guide for ethically sound practices by clinicians prescribing a drug with potential PGx variability, for the clinical researcher considering whether or not to study novel PGx markers in children, and for the investigator whose nontherapeutic PGx study may pose some risks?

In elaborating the concepts of benefit and risk, we consider several successful applications of PGx testing in pediatrics as well as historical cases in which the lack of information has led to harm. We also explore the ways in which the ethical issues involved in protecting a child's best interest go beyond concerns about physical well-being, drug dosing, or optimizing therapeutic benefit while minimizing bodily harm. Questions are currently being debated about how to appropriately obtain and manage the information that emerges through PGx testing in children.

There are many questions related to these concerns. For example, who should be responsible for ordering a pharmacogenetic genotyping panel that profiles responsiveness for a wide array of drugs such as those metabolized via the Cytochrome P450 isoenzymes? Given the wide variability in understanding of PGx testing and its relevance in a variety of situations, how should such information be transmitted to other clinicians and pharmacists who are caring for the child? When should genomic profiling be ordered, and when should it not be ordered? Are there situations in which a family requests testing and a clinician should refuse? If PGx testing is performed as part of a neonatal screening program, who should be the providers of such data? Should pediatricians be responsible for interpreting the data and facilitating patient management plans related to conditions that have an adult onset? At what age should children be informed of their specific liabilities? Should information be provided for conditions for which there is no known cure or treatment? If testing uncovers unexpected findings, what are the obligations regarding revealing this information? Is there ever harm that results from revealing information such as this? Who decides, and when? If a disease gene or disease susceptibility gene profile is discovered, should a subject be notified? Should the subject's family also be notified or is this a breach of confidentiality? [2].

It is important to note that these questions are also significant for adults and for the medical caregivers who treat them. This chapter focuses on children because of their vulnerability, but when adults are ill and their families are worried, many of the points that are made here are applicable: sick adults and their frightened families are not so different from sick children and their frightened families.

Pharmacogenomics raises other ethical questions that supersede issues centered on an individual patient and issues related to protection of information. For example, there is an important economic question: might pharmacogenomic criteria eventually become regulatory requirements for the development and testing of medicines? [3] Because the pharmaceutical industry functions on a business model, it is unlikely that it will develop medicines that benefit only a small, narrowly defined group of patients. Orphan drugs might therefore become more common. How should economic risk be weighed against the potential clinical benefit for a small group of patients? Furthermore, if pharmacogenomics information suggests only that a child is more or less likely to benefit from a drug, what threshold would exclude him or her from the use of that drug? Ten percent? Thirty percent? If a child is less likely to respond to a drug, is an insurance company obligated to pay for it? [4] If genetic variance peculiar to an ethnic group is found, what are the political, social, and economic implications? Might these be used to decide whether or not to develop certain medicines for certain populations? [5].

Finally, there are ethical issues related to social justice that are often uncomfortable and even contentious: in a time when tens of millions of people in the United States have limited access to basic healthcare, is the development of personalized medicine the most appropriate use of our resources as a society? Answers to questions such as these are influenced by changing realities accessible to study and description. If directing therapies to individuals based on their particular genetic (or other) features truly mitigates harm, this may provide long-term reduction in total healthcare costs. The application of PGx testing to patient care is already being facilitated by the increasing availability of genetic tests in clinical labs and relatively decreasing analytical costs.

On the other hand, if this technology cannot be meaningfully and broadly translated into the lives of individuals, or if harm outweighs benefit in a subset of the population, the priority of PGx testing might change. The question of whether or not PGx testing contributes to the societal goal of providing the best medical care at the lowest cost to the greatest number of people is an empirical one that can only be answered through careful investigation. The ethical question specifically related to testing in children is whether or not, and to what extent, children should be allowed (or asked) to participate in investigations or nonresearch testing while the final answers to the empirical questions are still uncertain.

The emergence of specific genetic variations, and their effects on drug therapy, has heightened pediatricians' awareness of the clinical application of PGx testing in the pediatric population, including both its potential benefits and its potential risks. Pediatricians also recognize important multifactorial challenges to widespread implementation of PGx testing in children, among them the following:

- Genetic factors are not the only influence on drug response; most children cannot provide informed consent to testing.
- The use of PGx test results to drive personalized medicine requires tests that are reliable and inexpensive and that provide fast turnaround time.
- The clinical relevance of genetic variations must be validated, and until this occurs it will be difficult to articulate the ways in which such testing can legitimately be viewed as in "the best interest" of a child.
- Clinical applications are currently limited to relatively few therapies.
- Health professionals need to receive education and training prior to integrating PGx testing, and this kind of training can be labor intensive and variably received by busy clinicians accustomed to the traditional approach to medical practice. There are unique ethical dilemmas associated with PGx testing and research in children, some of which may emerge only in the course of continued development of this technology.

When thinking about ethical issues related to children in any therapeutic or medical research setting, the primary goal is, and should be, to deepen our understanding of a child's changing ability to participate in ethical decision making, to be given the opportunity for meaningful assent or dissent where appropriate, to offer consent when developmentally mature or legally permitted, and to articulate the concepts of "best interest" in relation to a particular child. Just as medicine is "personalized" through these growing technologies, so moral deliberation regarding the appropriateness of pharmacogenomic testing should be "personalized" in the sense that decisions not only are grounded on moral consensus in the scientific and medical community, drawing on all that we have learned from past research efforts with children; they also align with the values of the families and children involved in studies.

41.2 HISTORICAL PERSPECTIVE

Medical research has involved children as subjects for hundreds of years [6]. Edward Jenner first used an experimental smallpox vaccine on his own 1-year-old son at the end of the eighteenth century. At the end of the nineteenth century, a 9-year-old child was the first human recipient of Louis Pasteur's rabies vaccine [7]. However, it was not until the first half of the twentieth century that regulations governing such experimental therapy began to be established and the importance of consent, assent, and parental permission became clearer.

Informed consent by human research subjects was first officially advocated in the Nuremberg Code, a code of ethics that was developed after the trials of German physicians who conducted morally outrageous human experiments in Nazi concentration camps during World War II (Table 41.1). Unfortunately, because this important code emphasized the absolute requirement of informed consent, it implicitly excluded children from participating in research on human subjects because they cannot give "informed consent" as framed by early codes of research ethics [8].

This issue was addressed more explicitly in 1964 when the World Medical Association adopted a set of research ethics principles now known as the Declaration of Helsinki [9,10]. These principles established the priority of human subjects' interests over those of science and society, and it sanctioned the participation of children in research if the child's responsible guardian gave permission.

After the publication of the Nuremberg Code, which implicitly excluded children from research, and before the Declaration of Helsinki, which explicitly allowed such research under specific conditions, experiments were nonetheless performed on children. Understanding this historical background illuminates the potential risks involved in any pediatric research, as well as those that emerge in nonresearch settings where testing or therapy is initiated without adequate evidential support. One of the most important of these was a series of radiation exposure experiments, in which children from the Fernald School in Framingham, Massachusetts, and researchers from the Massachusetts Institute of Technology participated, and which was funded by the Quaker Oats Company. In this experiment, children deemed mentally retarded were fed radioactive iron and calcium in their cereal [11]. President Clinton's advisory committee on human radiation experiments revealed the ethical consequences of inadequate parental permission in these experiments [12].

Another important set of experiments on children at around this time occurred at the Willowbrook institution (located on New York's Staten Island). These experiments were designed to follow the natural history of hepatitis in children also deemed mentally retarded [13,14]. Newly arrived children were infected with hepatitis virus by housing them with children already known to be infected. Participating children were kept in a special unit with better conditions and nutrition; if parents agreed to allow their children to participate, the children were admitted to Willowbrook more rapidly. The lead investigator did require

TABLE 41.1 Selected Guidelines on the Ethics of Biomedical Research with Human Subjects*

Guideline	Source	Year and Revisions
Fundamental		
Nuremberg Code [35]	Nuremberg Military Tribunal decision in *United States v Brandt*	1947
Declaration of Helsinki [36]	World Medical Association	1964, 1975, 1983, 1989, 1996
Belmont Report [37]	National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research	1979
International Ethical Guidelines for Biomedical Research Involving Human Subjects [38]	Council for International Organizations of Medical Sciences in collaboration with World Health Organization	Proposed in 1982; revised, 1983
Other		
45 CFR 46, Common Rule [8]	US Department of Health and Human Services (DHHS) and other US federal agencies	DHHS guidelines in 1981: Common Rule, 1991
Guidelines for Good Clinical Practice for Trials on Pharmaceutical Products [42]	World Health Organization	1995
Good Clinical Practice: Consolidated Guidance [44]	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use	1996
Convention on Human Rights and Biomedicine [43]	Council of Europe	1997
Guidelines and Recommendations for European Ethics Committees [45]	European Forum for Good Clinical Practice	1997
Medical Research Council Guidelines for Good Clinical Practice in Clinical Trials [46]	Medical Research Council, United Kingdom	1998
Guidelines for the Conduct of Health Research Involving Human Subjects in Uganda [47]	Uganda National Council for Science and Technology	1998
Ethical Conduct for Research Involving Humans [48]	Tri-Council Working Group, Canada	1998
National Statement on Ethical Conduct in Research Involving Humans [49]	National Health and Medical Research Council, Australia	1999

**CFR indicates Code of Federal Regulations. More extensive lists of international guidelines on human subjects research can be found in Brody [39] and Fluss [40]. An extensive summary of US guidelines can be found in Sugarman et al. [41].*

Source: From Emanuel, Wendler, and Grady [17].

a thorough consent process that included a two-week waiting period for full deliberation. But because Willowbrook was so crowded, critics subsequently argued that the expedited admission and special treatment amounted to coercion.

Both the Fernald and the Willowbrook cases underscored the need for ethical review and oversight of human subjects research involving children, even when parental permission is obtained.

During the early 1970s, Congress became increasingly concerned about research ethics, in part because of the syphilis study involving economically poor black men that was conducted by the US Public Health Service at the Tuskegee Institute. The investigation of this study led to the formation of the National Commission for the Protection of Research Subjects of Biomedical and Behavioral Research, which published the Belmont Report in 1979 [15]. This report embraced three principles that are now familiar and accepted as foundational for research involving human subjects: (1) respect for persons, (2) beneficence, and (3) justice.

The principle of beneficence grows out of the Hippocratic maxim "do no harm" and extends it to include maximizing the possible benefits and minimizing the possible harms to research subjects (minimizing harms is often formulated separately as a fourth principle, nonmaleficence) [16]. The principle of beneficence is operative in any discussion weighing the risks and benefits of human subjects research, especially when the subject is a minor. The principle of justice concerns the right and fair distribution of both the benefits and the burdens of research. This is an issue that has a particular urgency when clinical treatment and research are considered globally, and it is one that has become important as the ethics of pharmacogenomics mature. The concept of respect for persons includes two principles: that individuals be treated as autonomous agents and that those with less autonomy be entitled to protection. The latter principle is especially relevant in pediatrics and is somewhat fluid in its definition because of the growing autonomy of pediatric patients as they approach adulthood.

Because of the past abuse of research subjects who were harmed in studies that they did not understand or for which they had not given meaningful permission, it is understandable that "informed consent" has become the dominant concept in discussion about ethical research on human subjects. However, recent arguments have been put forth that informed consent is not always necessary for ethical clinical research, nor is it sufficient to qualify research as ethical. A recent review of major codes and declarations relating to human subjects research led to the proposal that there be seven requirements for ethical clinical research. These are worth keeping in mind both for pharmacogenomic research projects involving children and for the complex analysis of how to use such technology outside the research setting when evidence for efficacy is incomplete [17,18]. The requirements are these:

- The research must lead to enhancement of health or knowledge.
- The research must be methodologically rigorous.
- Selection of study sites and of individual subjects should be determined by scientific objectives and the potential for and distribution of risks and benefits.
- Given standard clinical practice and the research protocol, risks should be minimized, the potential for benefits should be enhanced, and risks must be outweighed by potential benefits to individuals and knowledge for society.
- Individuals should be informed about the research and should provide voluntary consent.
- Research subjects should have their privacy protected.
- Participants should have the opportunity to withdraw and have their well-being monitored.

These requirements are thorough and persuasive, but they must still be adapted to the conditions under which the research is conducted. Factors of health, economy, culture, religion, and technology affect how these requirements are translated into concrete, ethical action. For this reason, the specific context in which the research is conducted affects ethical deliberation.

In therapeutic and research efforts for children, ethical decision making comprises all of the considerations just discussed, in addition to the salient fact that the recipient of the risk or benefit is not generally the person who makes the decisions. A child may not be capable of assimilating sufficient information to meaningfully participate in an "informed choice." One need not enter the arena of international bioethics to find challenges to the role of "autonomy" in ethical decision making: such challenges are the daily fare in pediatric medicine and research because children's "autonomy" is an issue that changes from day to day.

41.3 ETHICAL SIGNIFICANCE OF BENEFIT, RISK, AND UNCERTAINTY IN PEDIATRIC PHARMACOGENOMICS

Many children are afflicted with potentially devastating diseases that are treated with medications that either can be, or will be, capable of being personalized with the aid of PGx testing. The history of pediatric medicine leaves little doubt about whether or not pediatric drug research is needed to improve medical care for children, and understanding this history can contribute to a better articulation of potential benefit to them—benefit that must always be weighed against the potential for harm. Accurate dosing of medications that carry risk is a compelling potential benefit, and the range of children for whom this benefit might be relevant is fairly broad: currently, there are ongoing studies to elucidate the relationship of pharmacogenetic variations and potential adverse drug reactions [19] for children with asthma, psychiatric and neurodevelopmental disorders, seizures, atopic dermatitis, human immunodeficiency virus (HIV), sickle cell disease, venous thrombosis, and acute lymphoblastic leukemia (ALL). Many of these conditions have an impact that goes beyond individual moral concern, extending to broader social and national concerns [20].

In the pediatric population, clinicians have often been forced to take a "trial and error" approach to using medications because of the lack of sufficient studies in children. Awareness of the potential harms involved in such an approach date back to 1938, when sulfanilamide elixir was prepared using diethylene glycol as the solvent. This led to 107 pediatric deaths, and there have been many pediatric drug misadventures since then. In the 1950s, widespread use of chloramphenicol resulted in grey baby syndrome, and sulfonamides were found to induce kernicterus. In the

1960s, tetracycline use led to bone and tooth deposition of the drug with functional and cosmetic consequences.

Two decades later, in the 1980s, it was discovered that additives in medications, rather than the drugs themselves, could cause adverse events. Heparin was thought to cause "gasping syndrome" in infants. However, it was discovered that accumulation of the preservative, benzyl alcohol, was to blame. Polysorbate found in intravenous Vitamin A products led to renal and hepatic failure and caused thrombocytopenia. An accumulation of propylene glycol, a solvent in many intravenous products administered to children, including multivitamins, lorazepam, and phenobarbital, caused iatrogenic hyperosmolality, metabolic acidosis, and clinical deterioration.

Understanding the behavior of medications in children, and the ways clinicians might use pharmacogenomic information to mitigate harm, requires participation by children in clinical trials. Children are vulnerable and developmentally dynamic, and they often lack the cognitive and conceptual ability to fully understand potential risks and participate in the decision making related to clinical care of research subjects. Therefore, parents, legal guardians, and healthcare professionals must be trusted to act in their best interests and make those important decisions for them. This is, of course, true for many areas in a child's daily life. For example, if we have to pick up a child at school and must take a younger sibling in the car as we do so, we are putting the younger sibling at risk for injury or death if we get into an accident, and we are doing so without any obvious direct benefit to the young child—we could have found a neighbor, a babysitter, or a relative to stay with the child if we had taken the effort to plan ahead. And yet we regularly make decisions that add risk to the lives of our children, and we take this in stride as part of life because such decisions are made in the context of a family's mutual goals and values. In biomedical research, where goals often arise from the broader priorities of the medical research community that is focusing on a more general notion of benefit that affects whole populations, questions about who takes the risk, and who receives the benefit, have a different status.

This vulnerability to risk without obvious direct benefit has meant that children are often excluded from clinical research in order to protect them. This in turn makes it harder to study the pediatric population in order to optimize therapies for it. There is an important and necessary tension between the desire to protect vulnerable children from exploitation in clinical research and the desire to benefit (often future) children through advances in our understanding of pediatric diseases and therapies.

In June 1996, the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) and the American Academy of Pediatrics (AAP) reviewed reports indicating that 10–20% of NIH-sponsored clinical research inappropriately excluded children. NICHD and AAP concluded that there is a tremendous need to include children in research. Indeed, balancing efforts to protect children from exploitation while working to advance discoveries that will benefit them is one of the most pressing and persistent parts of the ethical work that must pervade medical innovation in pediatrics.

From the time of the Nuremberg trials, the dangers of exploitation in medical research have been attended to and much harm averted. But with medical advances, we are also learning more about the real-world consequences in children's lives of being excluded from research. For example, asthma in children can truncate their ability to participate in many activities we associate with the joy of childhood, and in severe cases it can bring a child's life to an end. Can PGx testing help clinicians tailor therapy to maximize a child's ability to live fully despite asthma? Genetic variations in proteins acting as drug receptors can affect drug response in children.

Recent discoveries in asthma patients have shown that polymorphisms in the β-2 adrenergic receptor (coded by the *ADRB2* gene) may alter a child's response to β-2 receptor agonists. Several single-nucleotide polymorphisms (SNPs) have been identified that can heighten adverse effects or cause tachyphylaxis. Variants of the enzyme involved in leukotriene production, ALOX5, may affect response to the medications affecting these pathways, such as montelukast. A small subset of children with asthma may be poorly controlled because of steroid resistance due to decreased response to glucocorticoids and their anti-inflammatory effects related to genetic variations in cytokine production [21]. These insights are clinically relevant, and yet PGx data in children with asthma are scant.

As mentioned before, pediatricians often have to treat their patients in the absence of reliable data about the effects of the drugs they prescribe for them. This is reasonably construed as potentially harmful to children, so one relevant question that requires ongoing study is how the potential harm from this clinical practice compares to the potential harm of research trials or decisions made based on PGx testing in contexts where the meaning of the results is unclear. A child's response to pharmacotherapy is highly variable, and adverse drug events in some are often unpredictable. While certain children require a high dose of a medication to achieve a therapeutic response, other children with similar demographics have significant adverse drug reactions (ADRs) to the same medication in small doses.

The differences in response to drugs between children and adults are profound, and specific examples might serve to deepen ethical analysis of the nature of "benefit" and "harm" in pharmacogenomics. For example, in comparison to adults, infants and young children can develop acute central nervous system and hyperpyretic reactions to anticholinergic drugs, such as atropine and scopolamine, as well as life-threatening adverse reactions to valproic acid. Some drug response differences can be explained by pharmaco kinetic (PK) and pharmacodynamic (PD) variations in patient populations. However, studies have shown that children are at greater risk for severe ADRs than are adults, and this is a reasonable motivation to assess genetic variation and gene expression in order to make children's medications safer.

ADRs may be dose-dependent or dose-independent. Dose-related reactions can be predicted from the drug's known PK and/or PD. Certain populations, such as premature infants, are at high risk for dose-dependent ADRs because of the immaturity of their drug-metabolizing enzymes (DMEs). In the pediatric population, little is known about the PK and PD of the majority of medications used. Clinicians individualize drug therapy for infants and children by considering factors such as age, weight, drug allergies, organ function, concomitant medications, and historical drug response. The clinical drug response likely results from a combination of effects, including drug absorption, distribution, metabolism, and receptor site action, but it is possible that some medications are affected by specific DNA-sequencing variations. Human genes may affect the use of medications through several mechanisms, including drug-transporting proteins, metabolizing enzymes, and drug receptors. In children, the effects of ontogeny on PK and PD add an additional layer of complexity to PGx testing. For this reason, it is difficult to identify all causes of altered response in every patient. PGx testing provides additional information and a potential tool to guide appropriate therapy. The role of ontogeny and the effects of pharmacogenetic and developmental variations in DMEs, ion channels, transporters, receptors, and signaling pathways on the occurrence of ADRs in children require further exploration.

With the advent of pharmacogenomic analysis, it is possible to identify patients at risk of ADRs prior the administration of drugs. Cancer therapy is one very important area where PGx testing holds promise. For example, mercaptopurine is metabolized by the enzyme thiopurine methyltransferase (TPMT), which has a number of polymorphisms leading to the expression of three basic phenotypes. Approximately 89–94% of the population demonstrates "high activity" of this enzyme, while 6–11% display "intermediate activity." Importantly, a segment of the population, approximately 0.3%, displays "very low activity." In a single study of 147 patients receiving mercaptopurine for the treatment of ALL, homozygotes for low metabolic activity demonstrated severe hematologic toxicity despite a 50% reduction in dose. Prior to the understanding of the PGx testing results, this type of observation would have been idiotypic.

In addition to various mechanisms that can alter drug response, several families of drug-metabolizing enzymes have been identified that can have genetic polymorphisms with unique effects on PK/PD in children. These have potential for benefit, but also raise the important question (discussed later in the chapter) of implications of the connection between genetic polymorphisms and ethnic/racial background. For example, most Caucasians and half of African Americans do not functionally express the Cytochrome P450 enzyme 3A5. A commonly used benzodiazepine in pediatric patients is metabolized by CYP3A5 and CYP3A4. Polymorphisms of these enzymes result in decreased midazolam clearance and higher ADR risk. Another commonly used medication in the pediatric population is carbamazepine. There is a black-box warning from the US FDA about the potential of serious dermatologic adverse reactions, such as toxic epidermal necrolysis (TEN) and Stevens-Johnson syndrome (SJS) in those with the HLA-B*1502 allele of the HLA-B gene. This allele is found nearly exclusively in individuals of Asian descendent. The drug's labeling includes recommendations for testing for the HLA-B allele before initiation of use.

Codeine, a commonly prescribed opioid analgesic in children, recently created increased public awareness of variations in drug metabolism. Koren et al. published a case report in 2006 about an infant death believed to have been caused by morphine intoxication from breast-feeding [22]. The typical concentration of maternal codeine transferred from breast-feeding is 1–2 ng/mL. The authors noted that the infant's serum morphine concentration was 70 ng/mL. The mother was found to be heterozygous for the CYP2D6*2A allele and was an ultrarapid metabolizer of codeine. She was taking normal doses of acetaminophen with codeine for postpartum episiotomy pain, but this genetic polymorphism resulted in excessive production of morphine, which was distributed into her breast milk and transferred to the infant.

In August 2012, the FDA released a statement warning the public that a danger exists for children who are ultrarapid codeine metabolizers. It then conducted a safety review of its Adverse Event Reporting System (AERS) database from 1969 to May 2012. Ten deaths and three overdoses associated with codeine were found. These children were receiving doses of codeine following surgery, and many had a tonsillectomy or adenoidectomy. The doses of the opioid analgesic were within the typical dose range, and the signs of morphine overdose developed in 24–48 hours after consumption. The ages of the children ranged from 21 months to 9 years.

On February 20, 2013, the FDA added its strongest warning, a black box, to the drug label of codeine-containing products about the risk of codeine in postoperative pain management in children following tonsillectomy and/or adenoidectomy. Codeine is now contraindicated for that select patient population. Despite these warnings, the FDA does not recommend routine CYP2D6 genotype testing for use in this setting. The concern is that, although only up to 7 in every 100 individuals are CYP2D6 ultrarapid metabolizers, certain ethnic groups have been found to be more profoundly affected. Twenty-nine percent of North African and Ethiopian populations are ultrarapid metabolizers, whereas approximately 1% of Scandinavians, 10% of Spaniards and Greeks, and 6% of African Americans and Caucasians are affected. The only way to identify those who are at the highest risk for these adverse effects would be to perform PGx testing on children prior to administration of the drug.

Only a small number of all drugs and biological products marketed in the United States have had clinical trials performed in children. The majority of marketed drugs are not labeled for use in the pediatric population. Prior to the passage of the FDA Modernization Act (FDAMA) in 1997 and the Best Pharmaceuticals for Children Act (BPCA) in 2002 [23], fewer than 25% of drugs marketed in the United States were labeled for pediatric use [24]. While PGx testing provides an opportunity to optimize therapy, nearly half of all drug labeling still has insufficient information on safety, efficacy, or appropriate dosing for use in children.

Although there has been a slow decrease in "off-label" use of drugs in general pediatrics, neonatal information in labeling is even less available because neonates are a vulnerable subpopulation for whom end point development is lagging and in whom studies are more challenging. PGx testing can help explain or describe factors that influence drug absorption, drug metabolism, and drug disposition in children, especially given the many developmental factors in the early years of life that affect drug response, including organ development and protein levels [25]. But with the dynamic maturation of an infant's organ function, which has faster metabolism and more efficient clearance in the early years of life, it is arguably even more important to understand the effects that pharmacogenetic variations have in these children than in many other populations. The ethical challenge is this: Given that the potential benefits of PGx testing for a population, or a population subset, can be compellingly argued, and even though adults can consent to participate in clinical trials without obvious direct benefit to themselves, can children participate in studies that have more promise for benefiting future children than for benefiting the participating child?

41.3.1 What is "Consent" in Pediatrics?

The Nuremberg Code provides one starting point for understanding the meaning and importance of informed consent. The statement that "the voluntary consent of human subjects is absolutely essential" has been interpreted as legal capacity, the power of free choice based on knowledge and comprehension [26]. Consent, once seen as a single event, has come to be understood as more of a process [27]. It is not surprising, with the development of guidelines and goals for consent, that the process has become more intentional and more highly scrutinized.

The dominant theoretical framework for morally valid informed consent requires that four criteria be met: disclosure, understanding, voluntariness, and competence. Briefly, the core set of information that must be *disclosed* includes (1) the facts (such as risks, benefits, and alternatives) that patients (subjects) and providers believe relevant to the decision; (2) the recommendation of the professional; and (3) the purpose, nature, and limitations of consent. *Understanding* goes beyond disclosure because, while the elements disclosed can be objectively stated, true understanding involves many variables and is more difficult to assess.

Establishing and documenting understanding remains a great challenge, because information that has been disclosed but not understood contributes little to the ideal paradigm of informed consent. *Voluntariness*—another complex notion susceptible to interpretation—means at least that a decision is made without coercion or manipulation. Finally, *competence*, which is conventionally understood as the ability to perform a task, has also become a complex concept whose definitions derive from law, psychiatry, and philosophy. At base, a person's competence (or perhaps more accurately cognitive capacity) to make a particular decision relates to his or her ability to understand and think rationally about available choices, weigh the benefits and burdens and risks within the context of his or her life and values, and then use that understanding and logic to make a rational decision.

In the context of complex pediatric medical therapy or research, the process of consent can be influenced by the fact that parents are making high-stakes decisions on behalf of their own child. Even among clinicians, assessment of informed consent varies, and the experience of the clinician is relevant in it. Simon et al. found that clinicians with ten or fewer years of experience were more likely to say that the most important goal of informed consent is to explain the disease and treatment, and were more likely to suggest to parents that other children might benefit from the research [28]. The same study found that in the end, when reports from clinicians and parents are compared, clinicians are dissatisfied with aspects of consent that parents seem far more satisfied with. For example, clinicians expressed concern about information overload and about the fact that the consent discussion often occurs while the parents are still in a state of shock about the diagnosis.

Further insights into this difference have been gained from a number of studies addressing the perspectives of parents of children with cancer, a disease for which therapies are likely to be increasingly guided by technologies such as PGx. One study used three focus groups of parents of children with cancer to retrospectively examine their perceptions of the informed consent process [29]. High levels of stress were consistently reported, which were attributed to efforts to cope with multiple demands, including assimilating their child's diagnosis, nurturing and supporting the child, understanding the information given to them about diagnosis and treatment, getting to know an entirely new group of people involved in the child's care, and participating in the child's treatment.

Another important point was that the parents did not consistently distinguish research from their child's medical treatment. This finding underscores the importance of clearly explaining that research is optional during the process of "informing." If no distinction is made between research and treatment, the goal of informing has not been met. Physician-investigators often experience tension between their role as the patient's physician and their role as a researcher offering participation in a clinical trial designed to contribute to generalizable knowledge. This suggests that someone other than the physician-investigator should conduct the consent process. The study also suggested that the nature of clinical research, as well as the difference between research and proven current therapy, might be difficult for parents to grasp.

A second study in which parents of children with newly diagnosed cancer were interviewed supports these conclusions [30]. All of the participants recalled the diagnosis, and most (80%) recalled survival statistics. But only slightly more than half knew that the treatment protocol involved research and understood the concept of randomization. This finding is particularly important given that three-fourths of the parents thought that alternatives to enrollment in a randomized protocol had been insufficiently discussed. Because randomization is such an important tool for answering certain clinical questions of the sort that might involve PGx testing, this finding suggests that greater emphasis should be placed on explaining unfamiliar concepts during the process of "informing." Again, in this study the majority of participants were satisfied with the consent process.

A study by Kodish et al. confirmed the difficulty of effectively conveying the meaning of randomization during the consent process in pediatric leukemia trials [31]. In this multisite study, informed consent conferences were observed and audiotaped and were then compared to information acquired in interviews with parents shortly after the conference. The investigators found that while randomization was explained in 83% of cases, 50% of the parents did not understand it. Further, parents who did not understand randomization were more likely to consent to the randomized study than were those who did understand it, although this difference was not statistically significant ($p = 0.07$). These findings led to several recommendations for improving understanding, including a clear explanation of the differences between the randomized trial and off-study therapy, assessment of parental understanding of randomization, and further explanation when the idea was not yet grasped.

Other efforts have been made to improve the process of informed consent in pediatrics, especially in the case of complex clinical trials of the sort that increasingly incorporate PGx testing. For example, in a study assessing the possibility of staged consent for complex clinical research protocols, investigators had the option of obtaining consent over a 28-day period using a staged approach [32]. This option allowed parents and patients more time to discuss and absorb facts about the disease, the purpose of the trial, the design of the study, and the potential risks. Several measures in this study suggested that benefit accrues from the staged approach. There was greater understanding of treatment choices and of the distinction between a randomized controlled trial and standard therapy with the staged approach (80% understanding) than was found in the other studies (62.5%, $p = 0.05$).

Contemporary medicine that relies ever more on developing technologies such as PGx testing requires the introduction of new methods of informing patients and parents to meet the ethical imperative of informed consent in pediatric trials. Informed consent documents for clinical research trials are sometimes long and difficult to understand, with language that serves less to protect human subjects than to minimize institutional liability [33]. Therefore, some advances may take the form of technological and educational methods to explain ever more complex medicine and research to people not formally trained in a medical discipline, or even to busy medical professionals who cannot reasonably be expected to keep up with the explosion of information that is available. Potentially important are (1) the need for a "big picture" overview that would create a context for details, (2) delivery of information in small chunks, (3) use of multiple modes of information delivery, and (4) prioritization of information. Because of children's apparently natural affinity for interactive electronic devices, this approach may be helpful in addressing assent as well.

Concepts about the nature of consent in pediatrics are evolving with changes in contemporary medicine, including technologies that yield information the significance of which may not be certain. For example, some argue that informed consent may be too restrictive a concept and that "valid consent" should be substituted [34]. The three aspects of valid consent are personal competence (does the patient have the capacity to make the decision?), procedural competence (is the consent given correctly?), and material competence (is the procedure consented to appropriate for valid consent?). The concept of valid consent has been explored in the pediatric context [35]. Its value is that it emphasizes the patients' or parents' understanding of what is being consented to and recognizes that there are both rational and nonrational aspects to the decision-making process that must be understood.

The underlying concern is that informed consent has come to mean legally signed documentation rather than real understanding, insofar as that is possible. Because parents come from different backgrounds and because children of different ages differ in their ability to grasp complex concepts, the level of understanding that is attainable may vary from situation to situation. The notion of valid consent acknowledges this fact and counters the idea that a signed document is equivalent to informed consent.

The four elements of informed consent mentioned previously—disclosure, understanding, voluntariness, and competence—are useful for informed consent in adults, but informed consent in pediatrics is complicated by the fact that three parties are involved (parent, child, and clinician-investigator) and that the research subject is the child. In pediatrics, it is not autonomy (the basis of the four components of consent) that takes precedence but rather the *best interest* of the child [36]. The best-interests notion is clearer in a purely clinical setting than in a research setting, in which interventions are designed to contribute to general knowledge. Despite the difficulty of approximating truly informed consent in the setting of pediatrics, and especially pediatric research, the obligation to advance pediatric medicine lends urgency to continued efforts to offer the closest possible approximation to informed consent, which in pediatrics is a combination of parental permission and (as appropriate) the more complex concept of the assent of the child.

41.3.2 Assent in Pediatric Research

The ethics of assent is one of the most difficult issues faced in pediatric clinical trials, therapeutic interventions, and diagnostic or medical assessments such as PGx. Many adolescents, and even some younger children, possess the elements of competence or healthcare decision-making capacity (setting aside for the moment the legal definitions, which hinge on age rather than capacity). This is especially true, perhaps, of those who are exposed to long-term clinical trials and to the environment of a children's hospital for long periods of time. Certainly, younger children are not developmentally capable of comprehending complex protocols, but they do have some level of understanding that increases with age and experience [37]. How should assent be understood and when should it be required? But first, what do we actually mean by "assent" and how can it be distinguished from consent?

Consent refers to an active affirmative agreement to something by an autonomous agent (i.e., a person). It can be given by anyone who meets the minimal criteria of being in the state of mind and body to be an autonomous agent; generally most adults qualify as long as they have not been placed in a situation in which their decisions might be suspected to be the result of some form of coercion. The ability to give consent is also context and subject dependent in the sense that most people are generally not competent in all spheres of their lives: for instance, many individuals who are perfectly capable of making healthcare decisions or even agreeing to participate in research may not be competent to make (informed) decisions about automobile repair. For the purposes of this discussion, the relevant area in which cognitive capacity is important is in the domain of healthcare (and associated human subjects research).

In the United States, the legal age of consent is 18 years for most realms (nearly all jurisdictions, however, permit older adolescents who have not yet reached their majority to make some kinds of healthcare decisions—i.e., give consent), usually in regard to reproductive medical issues [38]. Assent is a more complex concept, and in general it describes a less active form of agreement to a proposed procedure or action than the more active and involved consent. It is employed in an attempt to engage children and adolescents in the process of consent that must be given by parents or other authorized surrogate decision makers [39,40]. However, what may qualify as acceptable assent (a grudging or sullen nod of the head, or a more energetic and participatory form of positive interaction) remains an area of reasonable concern.

Far less is known about assent than about consent. Much remains to be learned about the practices of institutional review boards (IRBs); the perceptions of parents, clinicians, and children about consent; and the methods that might be most effective in the assent process. Several studies have looked into these issues and documented the need for further work to address variability in IRB practices and the implementation of assent. One found that half of IRBs have a method they require investigators to follow when determining which children are capable of assent, while half have no such method; the majority rely on investigators' judgment [41]. IRBs need guidance on the implementation of requirements regarding assent.

A second study compared standards for assent as well as consent forms approved by 55 local IRBs reviewing three standardized multicenter research protocols [42]. Standards varied widely; 35 had separate forms and simplified language for assent, and 31 specified lower age ranges for obtaining assent in three studies. For a hypertension study, the age at which assent was required was 6–15 years. For a pain study, the age range was 6–12 years. And for a respiratory failure study, the range was 7–12 years. It is not clear why some IRBs consider a child of 6 capable of assent while others do not require assent until a child is 15 years old.

Do federal regulations help us understand assent? Most research on children is governed by subpart D of 45 CFR 46, which provides additional protections for children (as vulnerable subjects) beyond those specified in subpart A of this regulation. Subpart D, added in 1983, outlines three categories of research that a local IRB can approve: research not involving greater than minimal risk to participants (45 CFR 46.404), research involving greater than minimal risk but with the prospect of direct benefit to individual participants (45 CFR 46.405), and research involving no greater than a minor increase over minimal risk (with no prospect of direct benefit) but likely to provide generalizable

BOX 41.1 Principles Governing Children's Participation in Research Decisions

- Investigators should always respect children as persons. In particular, investigators, together with parents, should honor children's developing autonomy in decisions about research.
- Investigators should respect parents' roles in guiding their children's moral development and assessing their best interests. For example, parents should have discretion to determine the degree to which children should be encouraged to participate in activities that benefit others.
- Policies regarding assent, as well as IRBs' decisions with respect to particular protocols, should be sufficiently flexible to accommodate the wide range of medical, psychological, and contextual circumstances seen in pediatric oncology

Source: *From Joffe et al.* [44]

knowledge of the subject's condition or disorder that is vital to understanding or ameliorating it (45 CFR 46.406) (Box 41.1) [43].

Local IRBs cannot approve research that does not fall into one of these categories, and performance of the research is possible only if two conditions are met: (1) the local IRB finds that the research provides a reasonable opportunity to further the understanding, prevention, or alleviation of a serious problem affecting the health and welfare of children; and (2) the protocol is approved by the Secretary of Health and Human Services, after soliciting the opinions of an expert panel and providing for a period of public comment (45 CFR 46.407).

According to these regulations, assent means "a child's voluntary affirmative agreement to participate in research. Mere failure to object should not, absent affirmative agreement, be construed as assent" (45 CFR 46.402). If the IRB determines that the participants in certain categories of age and maturity are capable of providing assent, investigators must obtain it to proceed. When assent is required, a child's refusal is generally binding.

Can an IRB waive the requirement for assent? Yes, but only if (1) the intervention or procedure offers the possibility of direct benefit that is important to the health or well-being of the child, and if this intervention or procedure is available only in the research context (45 CFR 46.408(a)); or (2) the IRB determines that children below a certain age in a certain situation or with a certain condition have such a limited capacity to participate in the decision that they cannot reasonably be consulted.

The requirements are difficult to apply, especially in clinical trials for complex diseases in which there is uncertainty about the significance of aspects such as PGx testing. Cancer therapy is an excellent paradigm for considering issues such as this, in part because most of the children are enrolled in research studies in which the therapy is prolonged and carried out in the context of profound physical, spiritual, and psychological stress, and in part because cancer protocols are some of the most likely places where PGx testing will be used.

The Bioethics Committee of the Children's Oncology Group (COG) convened an international multidisciplinary task force to address assent in 2003. It identified a number of problems with the regulatory framework [44]. In brief, the regulations do not take adequate account of the dynamic moral and cognitive development of children and so allow a child either to have no formal role in decision making or to have the power of veto. Nor do the guidelines make clear what constitutes meaningful assent and so leave IRBs and investigators uncertain about when they have or have not hit the mark (though other groups have provided guidance in this area) [45,46].

Moreover, according to the COG task force, the regulations do not take into account the fact that some types of clinical research are more complex than other types, and so may be more accessible to the understanding of children. The regulations say that permission from the guardians and assent from the child are distinct decisions and they do not take into account the manner in which parents and children make decisions together. Finally, cultures in which autonomy is less emphasized are not accommodated by a regulation that potentially sets parental prerogatives against a child's veto power [47].

Considering these factors, the COG task force offered three principles governing children's participation in research decisions (see Table 41.2), from which a number of specific recommendations followed. The task force acknowledged that in some settings the principles will occasionally conflict. Therefore, the recommendations emphasize the importance of establishing procedures for the resolution of conflict rather than attempting to define universal rules. In that way, the absence of universal rules governing decisions in families with, perhaps, different cultural backgrounds and assumptions does not end conversation and negotiation. This approach is especially valuable in an increasingly pluralistic society and a global culture of information sharing. Much work remains to be done to understand the processes of communication and learning and the interactions that make up the permission-assent process. The flexibility of conflict resolution acknowledges that the obligation to obtain assent may change, depending on the time and situation, even for the same child. Developmentally appropriate ways to help the child understand the experience do not have the restrictions and limitations that accompany the binding aspects of assent.

For older adolescents, assent should be approached in the same way as for adult informed consent, even though

TABLE 41.2 Four Categories of Pediatric Research and Requirements for IRB Approval (45 CFR 46, Subpart D)

Level of Risk	Prospect of Direct Benefit	
	No	Yes
Minimal risk	Approval by IRB permitted(§46,404), conditioned on: • Permission from one parent/guardian • Assent of child, unless waived on capacity grounds	
Greater than minimal risk	Approval by IRB permitted(§46,406), conditioned on: • Minor increase over minimal risk • Research involves experiences that are commensurate with those inherent in children's actual or expected medical, dental, psychological, social or educational circumstances • Intervention or procedure is likely to yield generalizable knowledge about the subjects' disorder or condition that is of vital importance for understanding or ameliorating that condition or disorder • Permission from both parents if reasonably available • Assent of child, unless waived on capacity grounds	Approval by IRB permitted (§46,405), conditioned on: • Risks are justified by anticipated benefit to subjects • Relation of anticipated benefit to risk is at least as favorable as that presented by available alternatives • Permission from one parent/guardian • Assent of child, unless waived on capacity grounds or grounds of prospect of direct benefit that "is available only in the context of the research"
Not otherwise approvable, but presents opportunity to understand, prevent, or alleviate a serious problem affecting the health or welfare of children	Approval requires Department of Health and Human Services approval after consltation with an expert panel and opportunity for public review and comments (§46,407), conditioned on: • Research presents reasonable opportunity to further the understanding, prevention or alleviation of a serious problem affecting the health or welfare of children • Permission from both parents if reasonably available • Assent of child, unless waived on capacity grounds	

Source: From Joffe et al. [44].

parental permission may be required [48,49]. Some adolescents may have the capacity for consent but may choose to have their parents make decisions about their participation in research [50]. Several studies recently assessed assent in adolescents [51,52]. Our understanding of the differences between parents' and adolescents' perceptions of research is growing as we learn more about the assessment of risk (i.e., concern for physical safety) versus aversion (psychological discomfort) [53]. More than 3000 adolescents in the United States alone die of chronic illness and cancer each year. In such situations, when adolescents meet the criteria for functional competence, the broad ethical consensus is that decisional authority should be granted to them regardless of their legal decisional status [54]. However, this is an area in which there is wide divergence between what ethics may demand and what the law allows.

Most states and jurisdictions do *not* permit minors to be sole participants in medical decision making, including consenting to participate in research, unless they have been emancipated according to the laws of the state in which they reside. This restriction does not imply that adolescents should not fully participate in the decision-making process, and there certainly is no legal proscription against their doing so; it simply underlines the difference between ethical clinical (and research) practice and the legal framework. It might perhaps be better stated that in these situations, if the adolescent is going to be a research subject, assent is necessary but not sufficient.

41.3.3 Minimal Risk

One important ethical concept that guides discussion of pediatric research with no prospect of direct benefit is "minimal risk." With technologies such as PGx testing, articulating benefit can be complex because there is clear potential for harm if "trial and error" is used to assess best practices (as discussed earlier), and yet the significance of PGx testing in children, at this point, is often hypothetical or inadequately supported by evidence. The Belmont Report defines minimal risk as "the probability and magnitude of physical or psychological harm that is normally encountered in the daily lives, or in the routine medical or psychological examination, of healthy children."

The question that immediately arises is which children's lives should be the basis for deciding what is "normally encountered in daily life." These risks can vary widely, not only from country to country but within subgroups in a single country. Even if the notion of "harms or discomforts that average, healthy, normal children" may encounter" (as clarified by the National Bioethics Advisory Commission) can be roughly agreed on as a standard, the question remains whether it is ever acceptable for a parent to give permission for any risk greater than minimal risk in the absence of direct benefit to the child.

Because of concern that the "minimal risk" standard may exclude some pediatric research that is important and appropriate, the National Commission allowed parental discretion in giving permission for research that exposes children to a "minor increase" over minimal risk, given the fact that parents routinely permit their children to participate in activities, such as contact sports, that present more than a minimal risk, and given the potential benefits of such research to others. Therefore, a "minor increase" over minimal risk is permissible if the proposed research involves only procedures that such children might normally experience because of their medical condition, and if the proposed research has the potential to yield vitally important knowledge for understanding or ameliorating the child's particular condition.

The permissibility of a "minor increase" over minimal risk provides a conceptual way to avoid barriers to important pediatric research. But a difficult question remains: exactly what constitutes a "minor increase"? The term is used in the federal regulations, as previously noted, but it is not defined. Wendler and Emanuel proposed five possible standards for defining "minor" increase consistently among IRBs [55]. Their standards are (1) a fixed percentage above daily risks, (2) the confidence intervals around the risks of daily life, (3) the risks an ill child might encounter in an examination, (4) the "scrupulous parent" standard, in which minor risks are sufficiently similar to the child's experiences in daily life, and (5) the "socially acceptable" risk standard.

Wendler and Emanuel argue that the fifth standard is the most reasonable one. The "socially acceptable" risk standard takes as its starting point the fact that some children encounter greater risks in daily life than do others. Thus, it allows risks that are greater than those faced every day by healthy children, but it limits them to those that are socially acceptable. The socially acceptable risks that are greater than those experienced by the average healthy child may be more closely defined through research. In any case, this standard tethers the concept of "minor increase over minimal risk" to an evaluation that is nonarbitrary, that does not impede important beneficial research, and that does not justify greater risks for ill children simply because their daily lives are filled with more risk (because of their disease and its treatment) [56].

An alternative to the approach taken by Wendler and Emanuel has been suggested by Nelson and Ross. They argue that the "scrupulous parent" standard is ethically justifiable and that such an approach incorporates both minimal risk and a minor increase over minimal risk within a uniform standard that should be applicable to all children [57].

In PGx studies, as in other pediatric research, assent or dissent from a child with developing capacity can be overridden if the child's participation in the research involves a prospect of benefit to him or her such that the child's welfare is significantly jeopardized by failing to participate and the benefit in question cannot be obtained otherwise. The ethical tension arises when there is uncertainty about whether or not a child has the potential to benefit from a PGx trial. The weight of a child's assent or dissent is proportional to the uncertainty regarding benefit or harm. With relatively new technologies such as PGx, the potential clinical risks and benefits, as well as the social and legal ramifications, are still being tested; assessment and clear articulation of these risks and benefits thus become especially challenging.

Reducing harm in children while maximizing benefit is a primary driving force for advocates of PGx research and clinical testing in the pediatric population. The goal of PGx testing is ultimately to optimize therapy and to do so while decreasing unnecessary risk to children. Where it can be done without loss of benefit, risk reduction is a moral imperative, and so the ethical complexities related to PGx testing must be continually addressed as we learn more about its actual value to children.

41.4 FEDERAL REGULATIONS AND NATIONAL HEALTHCARE PRIORITIES RELATED TO PHARMACOGENOMICS

An active question in the healthcare community is whether or not advancing PGx testing into clinical practice should be a national healthcare priority. PGx research in children could lead to the formation of a new orphan disease group for which pharmaceutical companies fail to develop pharmacogenetic treatments for children because developing them for a minority among the population has no economic incentive. As a result, children would not benefit from the PGx developments. The sequencing of the human genome and the identification of links between specific genetic variants and diseases have created excitement over the potential of genomics to direct patient treatment toward more effective or less harmful interventions in all patients. However, many stakeholders disagree on whether our current understanding of genomic information is ready for clinical use. This is especially true in pediatric care because there have been limited studies and because there is worry that related technologies will add costs to the healthcare system without providing commensurate benefits. Others believe that healthcare costs could be reduced by identifying unnecessary or ineffective treatments. The answers to these concerns require careful study.

In order to address the lack of information available for pediatric dosing, the FDA began requesting data on pediatric medications on a voluntary basis in 1994. However, the number of new studies generated was small, and in 1997 Congress passed FDAMA, which included the Pediatric Exclusivity Provision, providing an incentive of an additional six months of patent protection in return for performing pediatric drug trials specified by the FDA. Given the success of the program, the BPCA of 2002 extended the economic incentives provided by pediatric exclusivity and the program was renewed in 2007.

The Pediatric Rule, which became effective in 1999, requires drug manufacturers to provide adequate labeling of their products in pediatric patients if the drug is likely to be used in this population. Although it was challenged in court in October 2002, the rule was codified by the Pediatric Research Equity Act (PREA), enacted in 2003, which also gave the FDA the authority to require pediatric studies of certain drugs. These regulatory changes have increased the number of drugs that undergo testing in children. By the late 2000s, still only 30% of all drugs used in the pediatric population had been studied in children. The lack of pediatric-specific drug indications and limited knowledge about PK/PD in children should facilitate further PGx research because it can be predictive prior to the first dose and offers the opportunity to improve our current strategies for dosing determination in children with a variety of pediatric diseases.

The FDA has provided guidance to aid in the translation of pharmacogenomic research findings into tools that can be applied to clinical care. It identified PGx as one of the areas of research that should be applied during drug development and subsequently launched its Critical Path Initiative in 2004. The following year, the FDA encouraged the voluntary submission of PGx data obtained during the drug development process in a document released to pharmaceutical manufacturers.

In 2008, the Pharmaceutical Research and Manufacturers of America (PhRMA) released its own white paper recommending that drug-metabolizing enzymes and transporter genotyping be incorporated into the drug development process. In September 2010, NICHD awarded a contract to establish a Pediatric Trials Network (PTN). PTN's primary objective is to provide an effective infrastructure for the ethical conduct of pediatric clinical trials for the BPCA drug development program, which include PK, safety, and efficacy studies. To continue its promotion of studies involving PGx, in February 2011 the FDA released a guidance in draft form entitled *Clinical Pharmacogenomics: Premarketing Evaluation in Early Phase Clinical Studies*, which provided recommendations on "when genomic information should be considered to address questions arising during drug development, and in some cases, during regulatory review." In early 2013, the FDA released an updated final guidance document in an attempt to help sponsors better conduct premarket evaluations in early-stage clinical studies and to emphasize that, although PGx approaches are unique to each product, the general principles of study design, data collection, and data analysis are broadly applicable. These recommendations will be necessary to promote PGx research in children.

Along with the FDA, other federal organizations, including the National Institutes of Health (NIH), provide funding for a collaborative research network dedicated to the discovery of genetic associations of drug response (the Pharmacogenomics Research Network, or PGRN). The members of the PGRN conducted research that helped to create the Pharmacogenomics Research Knowledge Base (PharmGKB), which has grown to become a publicly available PGx database.

In 2009, the Clinical Pharmacogenetics Implementation Consortium (CPIC) was formed as an extension of the PGRN and PharmGKB with a mission to assist in facilitating the development of clinical guidelines that can help providers interpret genetic testing results in order to appropriately adjust drug selection and dosing. As a result of this group's push for clinical relevance, the first guidelines were established in 2011, providing dosing recommendations based on TPMT genotype for the thiopurine drugs, which are commonly used in pediatric oncology patients. CPIC guidelines are intended to assist clinicians in understanding how available PGx test results should be used to optimize pharmacotherapy. However, they are not designed to guide clinicians in deciding whether or not to order PGx tests.

The recognition of specific genetic mutations that significantly influence drug response and/or toxicity is fueling an aggressive effort by many companies to develop suitable *in vitro* diagnostic tests. In order to quickly evaluate a child's genotype as a means of guiding optimal pharmacotherapy, diagnostic tests need to be readily available for use in the clinical setting, without the need to employ remote laboratories. There are several FDA-approved pharmacogenetic-designated diagnostic tests, including AmpliChip™, by Roche Diagnostics, and Invader® UGT1A1 Molecular Assay, by Third Wave technologies. PRO-PredictRx TPMT®, by Prometheus Laboratories, is available for screening TPMT alleles associated with increased immunosuppressive toxicity with azathioprine and 6-mercaptopurine.

A pharmacogenetic test for CYP2D6 is xTAG®, by Luminex Corporation. The CYP2D6 enzyme is involved in the metabolism of approximately one-quarter of all prescription medications, including opiates, selective reuptake inhibitors (SSRIs), tricyclic antidepressants (TCAs), neuroleptics, antiarrhythmics, and beta-blockers. There are also four FDA-approved warfarin pharmacogenetic test kits. These particulars are worth mentioning because they underscore one important tension in PGx testing: this is a technology that has potentially concrete applications, but while the data to support value are being generated, most third-party payers are reluctant to reimburse for testing because it is not currently considered a standard of care.

Whatever the potential benefits might be, multiple barriers preclude widespread implementation of PGx testing for personalized medicine as standard clinical practice across all medical fields. One of the major challenges for this industry is correctly identifying which of the estimated 30,000 genes are appropriate markers for influencing the physiologic response to drug therapy. It has been challenging to develop validated biomarkers and genetic assays, although the prospect of affordable exome and whole-genome sequencing will attenuate this barrier in the near future. Diagnostic genetic testing may be very useful for patient management and recurrence risk estimation as well as potential reduction. Twenty percent of the human genome has been protected under US patents in the past 31 years.

There is a global ethical concern related to patent holders not allowing wide availability of diagnostic care and access of information to high-risk patients. However, on June 13, 2013, the Supreme Court made a significant ruling against a molecular diagnostic company, Myriad Genetics, stating that human genes cannot be patented. From an ethical standpoint, the implications of genetic testing are likely to be far more profound with respect to disease assessment and/or prediction. The application of genetic information for the purposes of rational drug selection and dosing should be much less controversial. Will all individuals have equal access to pharmacogenetics and its benefits regardless of their socioeconomic status? Until a societal commitment to benefit sharing is actualized, market forces will continue to drive access to PGx testing. However, medical experts and legal advisors believe that this unanimous decision will have a lasting impact on genetic testing, inclusive of PGx testing, by making varieties more widely available and more affordable to all patients, including children.

41.5 TRANSLATION OF PHARMACOGENOMICS INTO THE CLINICAL SETTING

Most clinicians lack general knowledge of PGx testing and the interpretation of its results for clinical application. This may be one of the most pressing challenges that is preventing broad implementation of personalized medicine. Medical providers are in need of education as well as reliable and accessible information sources on PGx testing and clinical relevance. Medical and pharmacy schools are developing curriculum content in PGx. This will prepare future generations of providers, but there is little infrastructure in place to facilitate the education of clinicians already in practice [58]. In a recent survey of over 10,000 US physicians, 39% indicated that they obtain PGx testing information from FDA drug labeling [59].

More than 120 drugs have incorporated PGx information in their FDA labeling, but only a small number are currently used in pediatrics, such as atomoxetine, carbamazepine, codeine, fluoxetine, phenytoin, rifampin, valproic acid, mercaptopurine, and warfarin. FDA drug labels may contain information on genomic biomarkers such as gene variants, functional deficiencies, expression changes, and chromosomal abnormalities. They also may describe drug exposure and clinical response variability risk for adverse events, genotype-specific dosing, mechanisms of drug action, and polymorphic drug target and disposition genes.

Nevertheless, there is little labeling guidance on clinical application of results. Contradictory results obtained for certain variants have contributed to the lack of guidelines for clinical implementation. In the recent years, however, there have been more consistent results from investigators and an increasing need to establish working groups focusing on evidence-based clinical guidelines. One noteworthy example of PGx testing in pediatrics that illustrates its potential impact in children is the use of TPMT genotype and/or enzyme assay results in children with acute lymphoblastic leukemia (ALL) to predict chemotherapy sensitivity. The activity of TPMT, a key enzyme in the metabolic pathway of thiopurines, affects the safety and efficacy profile of mercaptopurine and thioguanine, two agents used in standard childhood ALL treatment. TPMT genetic polymorphisms can help explain the considerable differences in toxicities such as myelosuppression, which is experienced by children with TPMT variants that result in reduced TPMT expression and activity [60]. PGx testing can then be used in practice to predict and determine appropriate dosing, minimize toxicity, and optimize dosing, a practice which has been shown to impact outcomes for children with ALL [61].

The influence of TPMT activity on drug tolerability has led to changes in labeling for mercaptopurine, which now includes TPMT testing in dosing guidelines [62].

Other examples of PGx in children are Factor V Leiden testing prior to prescription of oral contraceptives for adolescents with a family history of thrombophilia or thrombosis, and occasionally VKORC1 and CYP2C9 testing in children starting warfarin [63]. However, PGx testing is far from universal in children: leukemia protocols usually only recommend TPMT testing for children who experience severe myelosuppression; thrombophilia testing in children is often controversial [64]; and PGx algorithms for warfarin dosing in children have not been validated [65].

Warfarin has also seen changes to its package labeling with regard to PGx testing. It is eliminated by hepatic metabolism to inactive metabolites; CYP2C9 is the primary responsible enzyme. CYP2C9*2 and CYP2C*3 are two allele forms that have been identified as decreasing metabolism. The package labeling indicates that the respective alleles occur in 11% and 7% of the Caucasian population.

A genetic polymorphism of the vitamin K metabolism enzyme (VKOR) has also been studied. Combined, these two polymorphic enzymes may account for 30–50% of warfarin dosing variability. Testing for VKOR and CYP2C9 is available, but there is no consensus on the interpretation and clinical application of the results. In order to enable clinical decisions, PGx guidelines, reference databases, and electronic platforms to integrate personal pharmacogenomic data into the clinical workflow would all be helpful in overcoming education barriers. If the potential for PGx testing is to be realized, and if miscommunication and faulty interpretation are to be avoided, innovative strategies must be found.

41.6 FUTURE IMPLICATIONS OF PHARMACOGENOMICS TESTING IN CHILDREN

Despite emerging data, the clinical implementation of PGx in pediatric patients has been a slow process since the completion of the Human Genome Project a decade ago. In addition to the ethical dilemmas unique to children, this gradual progression has led experts to further evaluate the ethical issues associated with future implications of PGx testing in children.

PGx testing can potentially place children in prognostic categories that have real consequences for their future, even into late adulthood [66]. This ultimately creates long-term privacy concerns and ethical issues related to adult-onset diseases revealed by PGx testing in early childhood, with the following possible consequences: discrimination related to the ability to afford PGx testing [67], race-based rather than individualized testing [68], lack of access to a drug to which an individual is not sensitive or that has a higher risk of complications [69], and the risk of test-based denials of disability, long-term care, and life insurance [70].

In light of these fears, it has been debated whether or not children should participate in testing that may reveal genetic predispositions to adult-onset disease or resistance to therapy if such findings will potentially impact a child's ability to obtain appropriate insurance coverage as they age. The refusal by parents or guardians to gain information that might affect the health of their child because of fear of discrimination based on genomic variability will hinder the advancement of PGx testing and research. Although PGx testing might hold promise for early discovery of disease and adaptation of therapy, the fear of no longer being insurable, at the point of discovery or at a later time in the child's life, might lead some parents and guardians to avoid testing that is potentially beneficial to a child, and ultimately lead to harm. As noted throughout this chapter, one of the most important areas in which PGx testing is relevant (and already actively used) is cancer therapy, and so the issue has received considerable attention from professional cancer organizations.

The American Society of Clinical Oncology (ASCO) has recommended that genetic testing for cancer predisposition be offered when (1) the individual has a personal history, family history, or features suggestive of genetic cancer susceptibility; (2) the test can be adequately interpreted, and (3) the result will aid in a diagnosis or in the medical or surgical management of a patient or family member who has hereditary risk for cancer [71]. ASCO has also recommended that the regulatory oversight of laboratories providing such tests be strengthened and that federal laws be established to prohibit discrimination by healthcare providers, employers, and insurance agencies on the basis of such information. Although some maintain that genetic discrimination will not affect insurance coverage, several surveys have shown that the public fears this possibility [72–74].

One important legislative step that has been taken to address such concerns is the Genetic Information Nondiscrimination Act (GINA) of 2008 (Box 41.2) [75]. This act was passed and went into effect in 2010. The intent of this legislation was to protect patients from discrimination by their insurance company or employer based on genetic testing results or family history. Another important issue is one of privacy and confidentiality and the possibility that genetic information obtained might be used to the detriment of a child sometime in the course of his or her life if legislative protections are weakened. The level of risk associated with such issues remains partially hypothetical, but an additional concern relates to informed consent with the potential for hidden future implications of this cutting-edge and evolving technology as well as for the legislation that will accompany it.

BOX 41.2 Genotypic Approaches to Therapy in Children

- Genetic information, not including information about the sex or age of any individual.
- An individual's genetic tests (including those done as part of a research study)
- Genetic tests of the individual's family members (defined as dependents and up to and including 4th-degree relatives)
- Genetic tests of any fetus of an individual or family member who is pregnant, and genetic tests of any embryo legally held by an individual or family member utilizing assisted reproductive technology
- The manifestation of a disease or disorder in family members (family history)
- Any request for, or receipt of, genetic services or participation in clinical research that includes genetic services (testing, counseling, education) by an individual or family member

Source: *From Ross et al.* [19]

In several ways, diagnostic genetic testing in children parallels PGx testing. Such testing is controversial and generally avoided until adolescence unless there is an immediate medical benefit identified, such as preventive measures to delay the onset of the disease, limit its severity, or prevent secondary morbidities. Also, early childhood diagnostic testing may be performed if there is benefit to another family member but no anticipated harm to the child [76]. However, data are needed before investigators, clinicians, and policy makers can conclude that genetic testing and pharmacogenetic testing in children are the same.

Like genetic testing, PGx testing provides ancillary information pertaining to the classification or prognosis of disease, predisposition to diseases or conditions, or information relevant to other classes of drugs for which the child is currently seeking treatment or manifesting symptoms. Ancillary information may also carry potential implications for family members when related to inherited polymorphisms. For example, returning to the case of therapy for the most common childhood cancer, testing for TPMT genotype prior to or during treatment with mercaptopurine for pediatric leukemia will provide information about the risk of myelosuppression and guide dose adjustments. Test results can provide prognostic information for the child and make available drug response information for a parent who needs to take mercaptopurine or azathioprine. Some PGx testing can have immediate implications for the child and at the same time have a bearing on health events that will occur much later in his or her life. Although less commonly prescribed in children, apolipoprotein E (APOE) genotype testing can guide warfarin dosing or statin selection, but it may also reveal the risk of an untreatable condition that the child may develop late in life, such as Alzheimer's disease [77].

The ethical dimensions of PGx warrant ongoing reassessment of the risks and benefits as more data are obtained [78]. With the exponential increase in health information sharing internationally, any robust development of ethical standards related to PGx will require not only national standards but also standards that cross the boundaries of nations engaged in such research. The World Health Organization's (WHO's) *Review of Ethical Issues in Medical Genetics* states: "All susceptibility testing should be voluntary, preceded by adequate information and based on informed consent…. Testing [of] children or adolescents should be carried out only if there are potential medical benefits to the child or adolescent or if an adolescent requests it for purposes of reproductive decision-making." [79] The American Academy of Pediatrics (AAP) and the American College of Medical Genetics (ACMG) recommend against predictive genetic testing for late-onset disorders unless there are specific interventions during childhood that will reduce morbidity or mortality [80].

The right to know or not know one's own PGx testing results is another controversial issue. Is it ethical for researchers not to reveal information that may benefit the child if consent rules prevent them from sharing incidental findings? Ethical debate continues as to whether investigators have a moral obligation to inform study participants about findings that are actionable and risks that can be mitigated. The ACMG recently recommended that incidental findings obtained in a clinical setting be disclosed to the patient. This statement excluded research studies.

When the consequences of the genetic disorder are treatable, the benefits of testing outweighs the negatives. This philosophy has driven pediatricians to support either earlier genetic testing or PGx testing or both. There is increasing interest in exploring the possibility of rapid whole-genome sequencing (WGS) for diagnosing genetic diseases in addition to newborn screening (NBS)—a mandated state-administered public health program designed to identify children with conditions that require immediate medical action to avoid death or severe disability [81]. NBS began in the 1960s with screening for phenylketonuria.

A federal panel is currently recommending screening for 31 core conditions, allowing states to screen for others as well. The evolution of NBS panels has advanced with advancing technology, and WGS as a new technology has the potential to dramatically expand the number of conditions screened. In this neonatal age group, monogenic disease progression can be frequent and fast, mandating an accelerated turnaround for molecular diagnoses so they can have an impact on clinical care decisions and potential therapies. Although there are over 3500 monogenic heterogeneous diseases that can result in profound morbidity and mortality, making the apparent benefit of testing very high, there are risks associated with PGx testing in infants. The hope of the WGS panel is that PGx results will also identify polymorphisms that may facilitate medication choices and appropriate doses. WGS is poised to eliminate the need for individual genetic tests.

WGS raises ethical issues about the creation of personal genetic databases holding information that individuals may or may not want to know and which may or may not remain privately secured. An enormous range of information can be gleaned from results that are highly accurate and actionable to genetic variations. However, these variations may not currently be associated with any known diseases, but they may create an information overload that will confuse clinicians until more robust studies can be conducted to elucidate such findings. In addition, as genetic testing and PGx testing knowledge advances, the future implications of WGS findings need to be considered. Unlike the NBS programs, there is no infrastructure currently in place to interpret the results of newborn sequencing. However, advocacy groups believe that some untreatable disorders should nonetheless have WGS information readily available so that parents can anticipate what lies ahead for their child. This delicate ethical balance is tipped when the driving force is to alleviate preventable harm in children in the immediate future while ensuring protection of their personal genomic data for future implications.

The storage and management of PGx data are ongoing issues. Most stored genetic data are de-identified. Unfortunately, however, the privacy of individuals cannot be fully secured, and the process of anonymization can be undone because some data is stored publicly online. This adds another complicating layer to the ethical decision that parents and guardians must make for their child, not knowing what PGx testing will reveal and how the information disclosed and the breaching of confidentiality will affect the child's future. The management and disclosure of ancillary clinical information must be evaluated, and recommendations need to be developed to guide PGx testing in children.

41.7 CONCLUSION

PGx testing in children has the potential to provide personalized medicine that will increase drug efficacy, decrease rates of toxicity, and improve overall outcomes in the pediatric population. The ethical challenges of childhood PGx testing and research are dynamic, so it is critical to reassess them as the level of the child's autonomy changes, as new PGx technologies are discovered, and as our knowledge of personalized medicine continues to expand.

REFERENCES

[1] Weinshilboum R, Wang L. Pharmacogenomics: bench to bedside. Nat Rev Drug Discov 2004;3(9):739–48.

[2] D'Ambrosio C, Gatta L, Bonini S. The future of microarray technology: networking the genome search. Allergy 2005;60(10):1219–26.

[3] Breckenridge A, Lindpaintner K, Lipton P, McLeod H, Rothstein M, Wallace H. Pharmacogenetics: ethical problems and solutions. Nat Rev Genet 2004;5(9):676–80.

[4] Service RF. Genetics and medicine. Recruiting genes, proteins for a revolution in diagnostics. Science 2003;300(5617):236–9.

[5] Lipton P. Pharmacogenetics: the ethical issues. Pharmacogenomics J 2003;3(1):14–16.

[6] Lederer SE. Children as guinea pigs: historical perspective. Account Res 2003;10(1):1–16.

[7] Fisher DJ. Resurgence of rabies. A historical perspective on rabies in children. Arch Pediatr Adolesc Med 1995;149(3):306–12.

[8] Diekema DS. Conducting ethical research in pediatrics: a brief historical overview and review of pediatric regulations. J Pediatr 2006;149(Suppl. 1):S3–S11.

[9] Levine RJ. International codes of research ethics: current controversies and the future. Indiana Law Rev 2002;35(2):557–67.

[10] Zion D, Gillam L, Loff B. The declaration of Helsinki, CIOMS and the ethics of research on vulnerable populations. Nat Med 2000;6(6):615–7.

[11] West D. Radiation experiments on children at the Fernald. Accountability Res 1998;6(1/2):103.

[12] Advisory Committee on Human Radiation Experiments. The human radiation experiments. Oxford University Press; 1996.

[13] Krugman S. The Willowbrook hepatitis studies revisited: ethical aspects. Rev Infect Dis 1986;8(1):157–62.

[14] Krugman S. Experiments at the Willowbrook State School. Lancet 1971;1(7706):966–7.

[15] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont report. United States Government Printing Office; 1979.

[16] Beauchamp TL, Childress JF. Principles of biomedical ethics. 5th ed. Oxford University Press; 2001.

[17] Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? JAMA. 2000;283:2701–11.

[18] Katz J. The Nuremberg code and the Nuremberg trial. A reappraisal. JAMA 1996;276(20):1662–6.

[19] Ross CJ, Carleton B, Warn DG, Stenton SB, Rassekh SR, Hayden MR. Genotypic approaches to therapy in children: a national active surveillance network (GATC) to study the pharmcogenomics of severe adverse reactions in children. Ann NY Acad Sci 2007;1110:177–92.

[20] Moran C, Thornburg CD, Barfield RC. Ethical considerations for pharmacogenomic testing in pediatric clinical care and research. Pharmacogenomics 2011;12(6):889–95.

[21] Tse SM, Tantisira K, Weiss ST. The pharmacogenetics and pharmacogenomics of asthma therapy. Pharmacogenomics J 2011;11(6):383–92.

[22] Koren G, Cairns J, Chitayat G, Leeder SJ. Pharmacogenetics of morphine poisoning in a breast fed neonate of a codeine-prescribed mother. Lancet 2006;368:704.

[23] Best Pharmaceuticals for Children Act, U.S.F.a.D. Administration, Editor. 2002. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/ucm049876.htm>.

[24] Choonara I. Unlicensed and off-label drug use in children: implications for safety. Expert Opin Drug Saf 2004;3(2):81–3.

[25] Yang YL, Lin DT, Chang SK, Lin SR, Lin SW, Chiou RJ, et al. Pharmacogenomic variations in treatment protocols for childhood acute lymphoblastic leukemia. Pediatr Blood Cancer 2010;54(2):206–11.

[26] Levine RJ. Ethics and regulation of clinical research. 2nd ed. Yale University Press; 1986.

[27] Katz J. The silent world of doctor and patient. Free Press; 1984.

[28] Simon C, Eder M, Raiz P, Zyzanski S, Pentz R, Kodish ED. Informed consent for pediatric leukemia research: clinician perspectives. Cancer 2001;92(3):691–700.

[29] Levi RB, Marsick R, Drotar D, Kodish ED. Diagnosis, disclosure, and informed consent: learning from parents of children with cancer. J Pediatr Hematol Oncol 2000;22(1):3–12.

[30] Kupst MJ, Patenaude AF, Walco GA, Sterling C. Clinical trials in pediatric cancer: parental perspectives on informed consent. J Pediatr Hematol Oncol 2003;25(10):787–90.

[31] Kodish E, Eder M, Noll RB, Ruccione K, Lange B, Angiolillo A, et al. Communication of randomization in childhood leukemia trials. JAMA 2004;291(4):470–5.

[32] Angiolillo AL, Simon C, Kodish E, Lange B, Noll RB, Ruccione K, et al. Staged informed consent for a randomized clinical trial in childhood leukemia: impact on the consent process. Pediatr Blood Cancer 2004;42(5):433–7.

[33] Grossman SA, Piantadosi S, Covahey C. Are informed consent forms that describe clinical oncology research protocols readable by most patients and their families? J Clin Oncol 1994;12(10):2211–5.

[34] Syse A. Norway: valid (as opposed to informed) consent. Lancet 2000;356(9238):1347–8.

[35] Spinetta JJ, Masera G, Jankovic M, Oppenheim D, Martins AG, Ben Arush MW, et al. Valid informed consent and participative decision-making in children with cancer and their parents: a report of the SIOP working committee on psychosocial issues in pediatric oncology. Med Pediatr Oncol 2003;40(4):244–6.

[36] Kodish E. Informed consent for pediatric research: is it really possible? J Pediatr 2003;142(2):89–90.

[37] Nelson RM, Reynolds WW. We should reject passive resignation in favor of requiring the assent of younger children for participation in nonbeneficial research. Am J Bioeth 2003;3(4):11–13.

[38] Coleman DL, Rosoff PM. The authority of “mature minors” to consent to general medical treatment. Pediatrics 2013;131(4):786–93.

[39] Bartholome WG. Informed consent, parental permission, and assent in pediatric practice. Pediatrics 1995;96(5 Pt 1):981–2.

[40] Wendler DS. 2006. Assent in paediatric research: theoretical and practical considerations. J Med Ethics 2006;32(4):229–34.

[41] Whittle A, Shah S, Wilfond B, Gensler G, Wendler D. Institutional review board practices regarding assent in pediatric research. Pediatrics 2004;113(6):1747–52.

[42] Kimberly MB, Hoehn KS, Feudtner C, Nelson RM, Schreiner M. Variation in standards of research compensation and child assent practices: a comparison of 69 institutional review board-approved informed permission and assent forms for 3 multicenter pediatric clinical trials. Pediatrics 2006;117(5):1706–11.

[43] U.S. Department of Health and Human Services. Federal policy for the protection of human subjects: additional protections for children involved as subjects in research (45 CFR 46, Subpart D). Federal Register. 1983.

[44] Joffe S, Fernandez CV, Pentz RD, Ungar DR, Mathew NA, Turner CW, et al. Involving children with cancer in decision-making about research participation. J Pediatr 2006;149(6):862–8.

[45] Committee on Bioethics. Informed consent, parental permission, and assent in pediatric practice. Committee on Bioethics, American Academy of Pediatrics. Pediatrics. 1995;95(2):314–17.

[46] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Research involving children: report and recommendations. Department of Health, Education, and Welfare, Publication (OS) 77–0004. 1977.

[47] Blackhall LJ, Murphy ST, Frank G, Michel V, Azen S. Ethnicity and attitudes toward patient autonomy. JAMA 1995;274(10):820–5.

[48] Rossi WC, Reynolds W, Nelson RM. Child assent and parental permission in pediatric research. Theor Med Bioeth 2003;24(2):131–48.

[49] Guidelines for the ethical conduct of studies to evaluate drugs in pediatric populations. Committee on Drugs, American Academy of Pediatrics. Pediatrics. 1995;95(2):286–294.

[50] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Research involving those institutionalized as mentally infirm. DHEW (OS) 78–0006. 1978.

[51] Olds RS. Informed-consent issues with adolescent health behavior research. Am J Health Behav 2003;27(Suppl. 3):S248–63.

[52] Lothen-Kline C, Howard DE, Hamburger EK, Worrell KD, Boekeloo BO. Truth and consequences: ethics, confidentiality, and disclosure in adolescent longitudinal prevention research. J Adolesc Health 2003;33(5):385–94.

[53] Brody JL, Scherer DG, Annett RD, Pearson-Bish M. Voluntary assent in biomedical research with adolescents: a comparison of parent and adolescent views. Ethics Behav 2003;13(1):79–95.

[54] Freyer DR. Care of the dying adolescent: special considerations. Pediatrics 2004;113(2):381–8.

[55] Wendler D, Emanuel EJ. What is a “minor” increase over minimal risk? J Pediatr 2005;147(5):575–8.

[56] Kopelman LM. When is risk minimal enough for children to be research subjects? In: Moskop JC, Kopelman LM, editors Children and health care: moral and social issues. Kluwer Academic Publishers; 1989.

[57] Nelson RM, Ross LF. In defense of a single standard of research risk for all children. J Pediatr 2005;147(5):565–6.

[58] Almarsdottir AB, Bjornsdottir I, Traulsen JM. A lay prescription for tailor-made drugs–focus group reflections on pharmacogenomics. Health Policy 2005;71(2):233–41.

[59] Stanek EJ, Sanders CL, Frueh FW. Physician awareness and utilization of food and drug administration (FDA)-approved labeling for pharmacogenomic testing information. J Personal Med 2013;3(2):111–23.

[60] Stanulla M, Schaeffeler E, Flohr T, Cario G, Schrauder A, Zimmermann M, et al. Thiopurine methyltransferase (TPMT) genotype and early treatment response to mercaptopurine in childhood acute lymphoblastic leukemia. JAMA 2005;293(12):1485–9.

[61] Relling MV, Hancock ML, Boyett JM, Pui CH, Evans WE. Prognostic importance of 6-mercaptopurine dose intensity in acute lymphoblastic leukemia. Blood 1999;93(9):2817–23.

[62] Relling MV, Hoffman JM. Should pharmacogenomic studies be required for new drug approval? Clin Pharmacol Ther 2007;81(3):425–8.

[63] Dietrich JE, Hertweck SP. Thrombophilias in adolescents: the past, present and future. Curr Opin Obstet Gynecol 2008;20(5):470–4.

[64] Thornburg CD, Dixon N, Paulyson-Nuñez K, Ortel T. Thrombophilia screening in asymptomatic children. Thromb Res 2008;121(5):597–604.

[65] Thornburg C. Anticoagulation in children: personalized strategies. Pediatr Health 2009;3(2):107–9.

[66] Barfield R, Kodish E. Pediatric ethics in the age of molecular medicine. Pediat Clin North Am 2006;53:639–48.

[67] Rogausch A, Brockmoller J, Himmel W. Pharmacogenetics in future medical care—implications for patients and physicians. Gesundheitswesen 2005;67(4):257–63.

[68] Bevan JL, Lynch JA, Dubriwny TN, Harris TM, Achter PJ, Reeder AL, et al. Informed lay preferences for delivery of racially varied pharmacogenomics. Genet Med 2003;5(5):393–9.

[69] Bates BR, Poirot K, Harris TM, Condit CM, Achter PJ. Evaluating direct-to-consumer marketing of race-based pharmacogenomics: a focus group study of public understandings of applied genomic medication. J Health Commun 2004;9(6):541–59.

[70] Dressler LG, Terry SF. How will GINA influence participation in pharmacogenomics research and clinical testing? Clin Pharmacol Ther 2009;86(5):472–5.

[71] American society of clinical oncology policy statement update: genetic testing for cancer susceptibility. J Clin Oncol 2003;21(12):2397–406.

[72] Thomas SM. Society and ethics -- the genetics of disease. Curr Opin Genet Dev 2004;14(3):287–91.

[73] Nowlan W. Human genetics. A rational view of insurance and genetic discrimination. Science 2002;297(5579):195–6.

[74] Rothenberg KH, Terry SF. Human genetics. Before it's too late–addressing fear of genetic information. Science 2002;297(5579):196–7.

[75] Genetic Information Nondiscrimination Act of 2008. <www.hhs.gov/ohrp/policy/gina.pdf> [accessed 28.04.13].

[76] Haga SB, Burke W. Pharmacogenetic testing: not as simple as it seems. Genet Med 2008;10(6):391–5.

[77] Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 1993;261(5123):921–3.

[78] Martinez W. Genetic testing of children and adolescents: ethical, legal and psychosocial implications. Princet J Bioeth 1998;1(1):65–75.

[79] Wertz DC, Fletcher JC, Berg K. Ethical, social and legal issues in pharmacogenomics. Pharmacogenomics J 2003;3(4):194–6.

[80] Nelson RM et al. Ethical issues with genetic testing in pediatrics. Pediatrics 2001;107(6):1451–5.

[81] Saunder CJ, Miller NA, Soden SE, Dinwiddie DL, Noll A, Alnadi NA, et al. Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. Sci Trans Med 2012;4(154):154. (3)

Chapter 42

PGx/Biomarker Utilization for Regulatory Decision Making

Yu Sakamoto,[a] Yasuto Otsubo,[a] Akihiro Ishiguro,[a] and Yoshiaki Uyama[a,b]
[a]*PMDA Omics Project, Pharmaceuticals & Medical Devices Agency (PMDA), Japan,* [b]*Department of Regulatory Science of Medicine, Graduate School of Medicine, Chiba University, Japan*

Chapter Outline

42.1 Introduction 951
42.2 Regulatory Activities and Guidelines Relating to PGx/Biomarker 951
  42.2.1 European Union: European Medicines Agency 953
  42.2.2 Japan: Pharmaceuticals and Medical Devices Agency 954
  42.2.3 United States Food and Drug Administration 955
42.3 Biomarker Qualification 956
  42.3.1 Significance 956
  42.3.2 Qualification Process 957
  42.3.3 Use of PGx/Biomarkers for a Specific Drug Product and Information in the Drug Label 958
42.4 Clinical Trial Design and Drug Development Utilizing PGx/Biomarker 960
  42.4.1 Designs of Clinical Trials with PGx/Biomarkers 960
  42.4.2 Study Design and Development Strategy for PGx/Biomarker Analysis 962
42.5 Future Challenges 963
Disclaimer 964
Glossary 964
Acronyms and Abbreviations 964
References 965

42.1 INTRODUCTION

PGx/biomarkers are among the most useful tools in drug development. They have substantially contributed to increased efficiency in drug development [1,2], which has resulted in major advances in personalized medicine as novel targeted drugs are being successfully developed for use in several therapeutic areas, including oncology. Because the application of PGx/biomarkers in drug development can reveal a drug's dose and target population more accurately than other conventional approaches, in the last decade more and more drugs have included PGx/biomarker information in their labels [3,4]. In this chapter, we describe the current regulatory guidelines on PGx/biomarkers and their qualification by regulatory agencies in different ICH regions. We also describe the design of clinical trials utilizing PGx/biomarkers for drug development, and we discuss the future challenges facing these new tools.

42.2 REGULATORY ACTIVITIES AND GUIDELINES RELATING TO PGx/BIOMARKER

One of the important roles of a regulatory agency is to evaluate relationships between the PGx/biomarker profile and the efficacy and/or safety of a drug. Thus, it is crucial to have guidelines detailing the points to consider in establishing evidence through the development of the drug. Since the early 2000s, the regulatory agencies of the European Union (European Medicines Agency; EMA), the United States (U.S. Food and Drug Administration; FDA), and Japan (Pharmaceuticals and Medical Devices Agency; PMDA) have worked to create frameworks for PGx/biomarker studies in drug development and have issued some guidelines (Table 42.1) [5]. In fact, two of them—ICH E15 and ICH E16—are harmonized guidelines prepared

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00042-6
© 2014 Elsevier Inc. All rights reserved.

TABLE 42.1 Regulatory Documents Related to PGx/Biomarker Studies in Drug Development

	EU	US	Japan
Terminology	Definitions for Genomic Biomarkers, Pharmacogenomics, Pharmacogenetics, Genomic Data and Sample Coding Categories (ICH E15) [6]		
General guideline on pharmacokinetics	• Guideline on the Investigation of Drug Interactions [8]	• Draft Guidance for Industry: Drug Interaction Studies—Study Design, Data Analysis, Implications for Dosing, and Labeling Recommendations [9]	• Guideline on Clinical Pharmacokinetic Studies of Pharmaceuticals [10] • Guideline on Methods of Drug Interaction Studies [11]
Sample collection and handling	• Reflection Paper on Pharmacogenomic Samples, Testing, and Data Handling [12]	• Guidance for Industry: Clinical Pharmacogenomics: Premarket Evaluation in Early-Phase Clinical Studies and Recommendations for Labeling [13]	• Guideline on General Principles for Clinical Trials Using Pharmacogenomics [14] • General Points to Consider in the Regulatory Submission of a Drug and Companion Diagnostics for Approval [15]
Assay (including companion diagnostics)	• Reflection Paper on Co-Development of Pharmacogenomic Biomarkers and Assays in the Context of Drug Development [16]	• Guidance for Industry and FDA Staff: Pharmacogenetic Tests and Genetic Tests for Heritable Markers [17] • Draft Guidance for Industry and FDA Staff: *In Vitro* Companion Diagnostic Devices [18]	• General Points to Consider in the Regulatory Submission of a Drug and Companion Diagnostics for Approval [15] • Technical Guidance on Development of *In Vitro* Companion Diagnostics and Corresponding Therapeutic Products [72]
Data submission	• Biomarkers Related to Drug or Biotechnology Product Development: Context, Structure, and Format of Qualification Submissions (ICH E16) [7]		
	• Guiding Principles: Processing Joint FDA EMEA Voluntary Genomic Data Submissions (VGDSs) within the Framework of the Confidentiality Arrangement [19]		
		• Guidance for Industry: Pharmacogenomic Data Submissions [20] • Guidance for Industry: Pharmacogenomic Data Submission – Companion Guidance [21]	
Clinical development and labeling (comprehensive documents)	• Reflection Paper on Methodological Issues associated with Pharmacogenomic Biomarkers in Relation to Clinical Development and Patient Selection [22] • Guideline on the Use of Pharmacogenetic Methodologies in the Pharmacokinetic Evaluation of Medicinal Products [23]	• Guidance for Industry: Clinical Pharmacogenomics: Premarket Evaluation in Early-Phase Clinical Studies and Recommendations for Labeling [13] • Draft Guidance for Industry: Clinical Pharmacology Section of Labeling for Human Prescription Drug and Biological Products—Content and Format [24]	
Clinical development in specific fields	• Reflection Paper on Pharmacogenomics in Oncology [25] • Reflection Paper on the Use of Genomics in Cardiovascular Clinical Trials [26]		• Partly described in the guidelines of clinical evaluation in each therapeutic area (e.g., anti-cancer drug [27])
Pharmacovigilance/ postmarketing activities	• Concept Paper on Key Aspects for the Use of Pharmacogenomic Methodologies in the Pharmacovigilance Evaluation of Medicinal Products [28]		• Request to Cooperate in research regarding severe adverse reactions (skin disorder, rhabdomyolysis and interstitial pneumonia) [29]

under the auspices of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). ICH E15, reached step 4 of the ICH procedure in 2007, and provides international definitions of PGx terminologies to facilitate the integration of pharmacogenomics and pharmacogenetics into global drug development and the regulatory approval process [6]. On the other hand ICH E16 reached step 4 of the ICH procedure in 2010, and aims to develop context, format, and structure for the regulatory submission of biomarker data to facilitate easy review and exchange of assessments among ICH members [7]. In this section, current key concepts on PGx/biomarker, as defined by each regulatory agency, are introduced based on published guidelines. It should be noted that we include our personal interpretations which may differ from the formal opinion of each agency.

42.2.1 European Union: European Medicines Agency

The EMA recognizes PGx as an important emerging technology that promises to optimize the development and use of medicines [30]. In 2000, a meeting on the use of pharmacogenetics in drug development was held for agency representatives and an informal group of pharmaceutical industry representatives [31]. It concluded with a demand for harmonized terminology and announced a need for pharmacogenetic clinical trial protocols and guidelines. In 2005, the PGx Working Party (PgWP) was formally established to provide the agency with expert support. The PgWP offers recommendations to the Committee for Medicinal Products for Human Use (CHMP) on all matters relating directly or indirectly to PGx.

In 2012, the EMA published a comprehensive final guideline with an aim to clarify requirements related to the use of pharmacogenetics in evaluating pharmacokinetic data [23]. The issues referred to in the guideline are the following: (1) situations and stages where pharmacogenetics-related pharmacokinetic studies should be performed; (2) *recommendations* and *requirements* regarding studies investigating the effects of polymorphisms at the ADME level; (3) evaluation of clinical impact and recommendations for further studies; (4) possible consequences of genetically determined differences in pharmacokinetic parameters for treatment recommendations and labeling; (5) special considerations related to drug–drug interactions (see also the general guideline for this issue [8]); and (6) effects of impaired or immature organ functions.

The guidelines highly recommend prospective DNA banking for genotyping throughout all phases of clinical development, even when there appear to be no obvious genetic influences on pharmacokinetics. This is in case unknown genetic factors are discovered in the later phase, including postmarketing.

According to this guideline, studies on the effects of pharmacogenetics on pharmacokinetics are *required* when the magnitude of the interindividual variation in drug exposure is so high that it may likely influence the safety and/or efficacy of the drug in genetically variable populations. Required situations are clearly discriminated from *recommended* situations using pharmacokinetic thresholds. If human *in vitro* studies suggest that more than 50% of a drug is cleared by a single polymorphic enzyme, the pathway can be considered important because the systemic exposure of the drug will be doubled, which is comparable to stepping up to a next stage in early phase I studies. In such situations, genotyping the participants of the first-in-human study is advised to avoid the safety issues associated with elevated exposure. Throughout phase I exploratory pharmacokinetic studies, expanded pharmacogenetic evaluation is recommended when more than 25% of a drug (comparable to the threshold for drug–drug interactions) is cleared by a polymorphic enzyme. In such situations, genotyping for the indicated genes in as many phase I studies as possible is required. As for active metabolites, genotyping is considered when a metabolite is estimated to contribute to more than 50% of the pharmacodynamic effect and a polymorphic enzyme is responsible for more than 25% of the *in vivo* formation or elimination of the substance.

If a significant difference in drug or metabolite exposure or distribution in a genotypically or phenotypically defined subset was indicated in earlier studies, genotyping, or alternatively phenotyping, of all participants in phase III trials is required. Care should be taken in the case of polymorphic transporters because intracellular or interorgan distribution of the drug may change, even when systemic exposure does not differ between different genotypes. The clinical consequences of the observed differences in drug exposure in genetic subpopulations depend on several factors. Unless and otherwise altered exposure to an active substance is of little consequence, dose adjustment should be considered to achieve an exposure in every genetic subset that is shown to be effective and safe, by genotype/phenotype-based or individual dosing.

In a reflection paper released for public consultation in 2011, genomic biomarkers were classified into three categories: predictive, prognostic, and surrogate [22]. The third category is not discussed in this chapter. Predictive biomarkers should have pretreatment characteristics that help to determine whether a particular subject is a good candidate for the treatment, with a test agent whose clinical utility is facilitated by pivotal trials conducted in defined patient populations. Although they can be indicators of efficacy and/or safety, it would be important to consider that more than one marker might be linked to a disease and might also influence drug response predictability. Prognostic biomarkers may correlate with disease outcome in either untreated or heterogeneously treated patients; therefore, they do not necessarily contribute to the clinical decision for treatment or intervention. The reflection paper suggests that exploratory studies for genomic biomarkers include a set of classifiers that translate biomarkers into a set of markers that predict clinical outcome. The paper also described two major purposes of genomic biomarkers. One is for patient selection, where the biomarkers can be used for: (1) better defining the disease and/or its prognosis; (2) excluding patients with increased risk; and (3) predicting drug response. The other is for determining whether the treatment algorithm can be used in clinical trials or clinical practice. As for using genomic biomarkers for selecting treatment strategies in drug development, the crucial step is to define and detail the analysis plan, including the criteria prospectively used for defining a positive response.

As the complexity of generating signals from a genomic biomarker is expected to be higher than that from other biomarkers, the reflection paper recommends that the following factors be carefully considered: consistency in sample collection; sample processing; assay methodology; and opportunities for misclassification. For genomic DNA variation and mRNA, careful evaluation of reproducibility is encouraged to avoid subject misclassification. In many cases, a considerable sample size is needed for identifying low-frequency signals from genomic biomarkers, which may have been generated in earlier exploratory studies as well as during later stages of drug development, including postmarketing. Sometimes prospective clinical trials, although intended to be carried out, turn out, for a number of reasons, not to be feasible. Therefore, the reflection paper suggests two possible alternatives: extrapolate previous scientific knowledge; and obtain data from retrospective samples or analyses. However, both alternatives have significant limitations. In order for primarily retrospective studies to be credible, certain requirements must be met. Besides confounders, selection bias and measurement bias are important in genomic biomarkers development. The former can be minimized by properly designing the prospective study, but on the other hand, could impact retrospective analyses. The latter might affect retrospective analyses, especially meta-analyses with different instruments or methodologies. In addition, two distinct multiplicity problems can arise during genomic biomarker development: (1) the use of multiple markers to identify the one that has a sufficiently strong association with the outcome; (2) the issue of multiple testing during the clinical trial. The reflection paper points out the importance of characterizing the diagnostic performance of the biomarkers, as well as the factors that would affect their performance. The clinical validity of a marker is said to be a complex interplay of sensitivity and specificity and the penetrance of the genomic abnormality or mutation. For novel or specific genomic biomarkers, it would be necessary to develop specific assays as companion diagnostics (CDx) in parallel with drug development. If a particular assay specific to a genomic biomarker is employed in confirmatory studies, it may be necessary to link it to the marker's value.

The EMA provides current views on regulatory matters regarding pharmacovigilance. A concept paper was published in 2011 to address the scarcity of the PGx/biomarker information on drug effect during postmarketing [28]. It is anticipated that recognition of a genetic variant may permit estimation of the likelihood of adverse drug reactions (ADRs) or the effectiveness of genetic information or tests. It is expected that a draft guideline will be developed soon.

The agency has also published several other specified documents: reflection papers on PGx in cardiovascular clinical trials and oncology; on samples, testing, and data handling; and on co-development of certain biomarkers and their assay platforms [12,16,25,26].

42.2.2 Japan: Pharmaceuticals and Medical Devices Agency

The Pharmaceuticals and Medical Devices Agency (PMDA) has worked together with the Ministry of Health, Labor and Welfare (MHLW) to facilitate drug development utilizing PGx. In 2005, a notification was released to gather voluntary information from industries in order to gauge the development situation at that time [32]. Three years after the notice was circulated, a guideline on "General Principles for Clinical Trials Using pharmacogenomics" was published in 2008 [14]. Since the PGx in drug development would enhance drug safety and/or efficacy, increasing numbers of clinical studies are expected to include PGx analyses. The guideline mainly focuses on points to consider when collecting genomic samples, both pre- and postmarketing, for clinical trials. It describes three situations where genomic sampling is planned: (1) both target genes and time of analyses are predefined; (2) target genes are predefined but the timing of future analysis is not; and (3) neither the target genes nor the timing of analysis are predefined (note that the third situation may include analyses not directly related to a investigational drug). For all three categories, clarification of the purpose of the planned analyses and the available knowledge on the predefined or putative target genes in study protocols is required. Above all, written informed consent from participants and appropriate sample anonymization is necessary for any purpose. In addition, it has been recommended that clarification of the key information from pharmacogenomic analyses (e.g., analytic subset, sample handling including anonymizing method, and disclosure policy) in study protocols or other operational procedures is clearly documented.

Two guidelines, one on "Clinical Pharmacokinetic Studies of Pharmaceuticals" [10] and the other on "Methods of Drug Interaction Study" [11], were published in 2001. The guideline for drug interaction study is currently under revision to include the latest scientific knowledge, including PGx/biomarker, and will be available in the near future. In response to the globalization of drug development, a guideline entitled “Basic Principles on Global Clinical Trials” [33] and an accompanying document entitled “Basic Principles on Global Clinical Trials (Reference Cases)” [34] were published recently. These recommend performing PGx/biomarker analysis, including genetic tests using samples collected in global clinical trials.

In the PMDA, a “pharmacogenomics discussion group” was originally established in 2005 and was reorganized as “PMDA Omics Project (POP)” in 2009 to cover PGx and other omics issues, including proteomics and metabolomics. POP has worked to promote scientific discussion on PGx/biomarker and PGx/biomarker-guided drug development. Since 2009, a new scientific consultation service focusing on PGx/biomarker has been started to encourage discussions between the industry and the PMDA on the use of PGx/biomarker in drug development and regulatory submission. In 2012, the consultation service was reinforced by a new category: planned stage and follow-up consultations (for more details see Section 42.3).

In 2012, the “Companion Diagnostics Project Team” was established in the PMDA to promote co-development of a companion diagnostic with a drug. One achievement of this project is a guideline entitled “General Points to Consider in the Regulatory Submission of a Drug and Companion Diagnostics for Approval” published on December 26, 2013 [15]. Afterwards, more practical guidance for developing companion diagnostics entitled “Technical Guidance on Development of *In Vitro* Companion Diagnostics and Corresponding Therapeutic Products” was published [72].

In terms of PGx/biomarker use in the postmarket stage, a notification entitled “Request to Cooperate in Research Regarding Severe Adverse Reactions” [29] was published by the MHLW. In this research, relationships between the genetic profiles and severe adverse reactions, such as Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN), rhabdomyolysis, and interstitial pneumonia, have been investigated. In this regard, it should be noted that cooperation between industry, hospital, and research institute is the key to facilitating research by effectively acquiring samples from patients who suffer from ADRs. Outcomes of this research have been published in several scientific journals [35–37].

42.2.3 United States Food and Drug Administration

In 2004, the Food and Drug Administration (FDA) released a white paper addressing the significant decrease in applications submitted on new drugs and biologics [38]. In this paper, the FDA regarded biomarkers as a powerful scientific and technical tool for improving the predictability and efficiency of medical product development. The FDA Genomics Group in the Office of Clinical Pharmacology works to advance the application of genomics for the benefit of patients and society. Genomics is recognized as an integral discipline in the discovery, development, regulation, and use of medications.

In 2005, to facilitate scientific progress in PGx and the use of pharmacogenomic data in drug development, the FDA published a guidance for genomic data submission [20]. This guidance, with the help of some decision trees, describes when to submit pharmacogenomic data during the development and review process, what format and content should be provided, and how and when the data are to be used in regulatory decision making. The guidance also makes a distinction between *valid* biomarkers, which may be appropriate for regulatory decision making, and either *observational* or *exploratory* ones, which alone are insufficient. Furthermore, the guidance proposes an additional distinction between *known* valid biomarkers and *probable* ones, the former “[having] been accepted in the broad scientific community” and the latter “[appearing] to have predictive value for clinical outcomes but may not yet be widely accepted or independently verified by other investigators or institutions.” Data on probable valid biomarkers are not necessarily required to be submitted to the Investigational New Drug (IND), but its submission is recommended for unapproved NDAs or BLAs. At the research stage of establishing the validity of a biomarker, pharmacogenomic data submissions are not required, but the FDA has been encouraging voluntary genomic data submissions (VGDSs) following the establishment of the FDA-EMA joint VGDS process [19]. Currently, this process is recognized for voluntary exploratory data submissions (VXDS). Any data submitted as VXDS will be reviewed for scientific and informational purposes, and will not be used for regulatory decision making. Based on review experience gained from the VGDS process, INDs, NDAs, and BLAs, a companion draft guidance mainly for data collection and handling was released in 2007 [21].

In 2013, the FDA published a guidance document for clinical evaluation entitled “Guidance for Industry: Clinical Pharmacogenomics: Premarket Evaluation in Early-Phase Clinical Studies and Recommendations for Labeling” [13]. This document focuses on when and how to evaluate, in the early stages of drug development, the contributions of pharmacogenomic variations to the individual differences observed in pharmacokinetics, pharmacodynamics, and safety or efficacy. Three key purposes of pharmacogenomic assessment in early-phase clinical trials are proposed in the guidance to identify: (1) populations that should receive modified dosage; (2) responder populations; and (3) high-risk groups. The genomic influences suggested in the early-phase or confirmatory trials may include subject enrichment according to genomic information, dosage based on genomic variability, or a priori primary end point testing in a genomically defined group. The FDA also emphasizes that baseline DNA sampling should be performed on all subjects employed in the clinical trial to eliminate any potential bias associated with incomplete sampling. Also suggested is storing of all samples or appropriate data for at least 15 years in case of later need, as otherwise any known genetic or other predefined potential factor will not be available at that time.

Preclinically identified pharmacogenomic factors should be taken into consideration for the design and analysis of clinical trials. When an exposure is variable among patients but correlated with drug response, the results may be considered for dose selection at a later stage. Use of approaches involving candidate genes is proposed when the factor that influences the pharmacology and pharmacokinetics of the drug is well known. High-throughput platforms can also be utilized, particularly when the pharmacologic property of the drug is not clear; however, in such cases it is important to experimentally confirm the physiological roles of the newly identified factors. To avoid misclassification as a result of not testing, efforts should be made to include rarer alleles of the factor of interest.

The FDA guidance proposes strategies for designing three types of clinical pharmacogenomic studies: PK/PD studies in healthy volunteers; PK/PD studies in patients; and dose–response studies. In general, collection of DNA samples from all healthy volunteers who participate in the single- and multiple-dose PK studies is encouraged, while taking notes on the drugs that convert inactive compounds into active metabolites through polymorphic pathways. In cases where important variability is found in PK/PD studies involving healthy volunteers, its significance should be considered when designing these studies on patients. The findings from such studies can increase the average effect, decrease toxicity, and increase the possibility of overall study success. Genomic factors should be taken into account for dose–response studies if any previous PK and/or PD study suggested genomic influence on systemic exposure–response, or on efficacy and safety responses. Genomic information obtained from such trials can be integrated into later analyses of dose–response or concentration–response models. The guidance also suggests specific considerations while designing studies. For example, to compare PKs in genomic subgroups, studies can be designed like those for evaluating the effects of hepatic or renal impairment. For PD, study designs are less well developed because, while genomic markers are usually explored in early-phase studies, clinical outcomes are evaluated later in drug development. The relationship between a genomic marker and PD suggested in later trials often needs to be prospectively confirmed with respect to genetic subsets.

The FDA guidance describes cases where patients with certain genotypes might be excluded: when the patients would experience toxicity; would not show response to the drug due to low exposure; and lack a drug target. However, it is cautioned that many quantitative markers do not lead to binary response, so early-phase trials should include subjects with a range of target levels in order to explore optimal cut-offs. In early-phase exploratory and observational trials, profiles of pharmacogenomic markers are often examined using nonrandomized designs, involving a small population, that are not adequately powered. Thus, the guideline provides five statistical concepts that are most relevant to clinical PGx: (1) avoid confounding factors; (2) control multiplicity; (3) characterize marker performance quantitatively; (4) address the potential of incorrect population stratification; and (5) assess the reproducibility of the genetic finding. Pooled analyses of trials with similar protocols are proposed, and replication of the findings or experimental studies are recommended. Any retrospectively defined relationship should be replicated to avoid false positive and potential bias. Using a case-control approach to investigate pharmacogenomic markers related to ADRs would lay weight not only on high sample acquisition rates, but also on a clinically meaningful definition of cases and clearly described controls. At the end of the guidance, the FDA states that drug labels should generally include information on PGx only if it is useful to inform prescribers about the impact of genotype on phenotype, or to indicate whether a genomic test is available.

The efforts described above clearly indicate that EU, U.S., and Japanese regulatory agencies work with similar concepts to promote PGx/biomarkers in drug development to achieve personalized medicine in clinical practice. However, their thinking is still different in some respects, such as regarding sample collection, qualification and the requirements for PGx-related PK studies. As a next step, more international cooperation is necessary to harmonize the PGx/biomarker guidelines because drug development has been radically globalized [39–41] and the same PGx/biomarker data may have been submitted to multiple regulatory agencies (see Section 42.5).

42.3 BIOMARKER QUALIFICATION

42.3.1 Significance

To obtain good PGx/biomarker-related scientific evidence in a study, it is very important to use a PGx/biomarker qualified by a regulatory agency for its proper use in the drug development [42]. Use of an unqualified and unacceptable biomarker may result in misinterpretation of the acquired data because of false signals (false-positive or false-negative), and therefore cannot be used in regulatory decision making. "Qualification" means that, within the stated context of use, the results of assessment with a biomarker can be relied upon to adequately reflect or forecast a biological process, response, or event, and to support its use in new drug development, ranging from discovery through post-approval [7]. A qualified biomarker not only has clinical use but also has nonclinical use. To qualify a biomarker, it is important to clearly define its "context of use" according to the ICH E16 guideline [7,42]. A biomarker's context of use with respect

TABLE 42.2 Comparison of the Qualification Processes of Three Regulatory Agencies

	EMA	FDA	PMDA
Guidance document or reference website	• Qualification of Novel Methodologies for Drug Development: Guidance for Applicants • http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000319.jsp&mid=WC0b01ac0580022bb0	• Qualification Process for Drug Development Tools (Draft guidance) • http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284076.htm	• Special scientific consultation regarding PGx/biomarker qualification • (http://www.pmda.go.jp/operations/shonin/info/consult/file/0302070-betten4.pdf, in Japanese only)
Scope	• Innovative drug development methods and tools, including imaging and drug/diagnostic co-development.	• Any biomarker for clinical or nonclinical context of use, including imaging but not drug/diagnostic co-development.	• Any biomarker for clinical or nonclinical context of use, including imaging but not individual drug/diagnostic development such as drug/diagnostic co-development.
Option	• Qualification advice • Qualification opinion	• Not specified	• Consultations for planning a study • Consultations for qualifying PGx/biomarker data
Fee	• Full scientific advice: €75,500 • Follow-up advice: €37,700 • 90% reduced rate for small and medium enterprises	• Free	For evaluating study data • First consultation: ¥3,114,900 • Follow-up: ¥948,300 For a study plan and strategy • First consultation: ¥1,142,800 • Follow-up: ¥414,600
Output	• Formal regulatory opinion	• Formal regulatory opinion	• Formal qualification opinion
Standard timeline (from application to complete report)	• 190 days	• Not specified	• About 24 weeks

to qualification can be narrow or broad. Depending on the context of use, it might be used for several drugs within one class of drugs, or even for several classes of drugs.

42.3.2 Qualification Process

Prior to being widely used in drug development, a biomarker should be qualified by a regulatory agency in accordance with its intended objective and context of use. To confirm acceptability, each regulatory agency—EMA, FDA, and PMDA—has a biomarker qualification process as summarized in Table 42.2. Major points for qualification according to PMDA are as follows:

- The data submitted for qualifying a biomarker should be prepared according to the format and structure described in the ICH E16 guideline.
- The applicant should clearly define the "context of use of the biomarker" according to the ICH E16 guideline. The term "context of use" is a clear statement that describes the manner of use, interpretation, and purpose of use of a biomarker in drug development. The elements describing context of use should include (1) general area; (2) specific biomarker use; and (3) critical parameters defining when and how the biomarker should be used.
- A meeting is to be held having two main objectives: (1) to discuss the study design and strategy for qualifying a PGx/biomarker; or (2) to evaluate biomarker data for regulatory qualification.
- The applicant is required to answer all inquiries from the PMDA.
- Face-to-face meetings between the PMDA and the applicant will be held as needed.
- The qualification meeting covers PGx/biomarkers that are not directly linked to specific products; such biomarkers should be discussed during the PMDA standard scientific consultation process.
- The PMDA will provide an official consultation report to the applicant.

The concept of and required documents for the qualification meetings held by the EMA and the FDA are similar to those of the PMDA, although there are some differences, such as in fee, timeline, and implementation status (formal/informal). The ICH E16 guideline is expected to facilitate data submission for biomarker qualification on a global basis. An internationally qualified biomarker is one of the most useful tools for global drug development. In particular, PGx/biomarkers used in global clinical trials should be acceptable to multiple regulatory agencies, including those of the country where the trial is conducted; if not, the collected data cannot be used for regulatory purposes [39]. Therefore, an applicant is encouraged to seek parallel advice from EMA, FDA, and PMDA regarding biomarker qualification, although there is no formally established process for qualifying a biomarker across all three agencies at present.

42.3.2.1 Case Example of Qualification

Table 42.3 lists some of the biomarkers that have been qualified by EMA, FDA, and/or PMDA. Seven urinary biomarkers, such as KIM-1, cystatin-C, and beta2-microglobulin, have been qualified by these agencies for nonclinical use to detect nephrotoxicity in rats [43,44]. This means that the seven biomarkers, in the specific context "to detect drug-induced acute urinary tubular changes or acute glomerular changes/injury in rat GLP studies," are now qualified at the same level by all ICH agencies. Thus, they can be used for regulatory submissions during the drug development process. Moreover, EMA and FDA have qualified urinary clusterin for an additional context of use and a new urinary marker, renal papillary antigen-1 (RPA-1), for detecting acute drug-induced renal tubule alterations. The FDA also concluded that serum/plasma cardiac troponins T and I (cTnT and cTnI) can be used as qualified biomarkers for studies assessing potential cardiotoxicity in rats and dogs.

At present, no clinical biomarker is qualified by all three agencies. So far, only EMA has issued an opinion on the use of clinical biomarkers—namely, tau and Aβ42 in cerebrospinal fluid—for Alzheimer's disease [45]. Hippocampal volume (atrophy) and PET amyloid imaging have been qualified for Alzheimer's disease [46,47]. To promote the use of such biomarkers in clinical trials, especially global trials, it is necessary that the biomarker be qualified by not one agency but multiple regulatory agencies for acceptance as internationally qualified.

To increase the number of qualified biomarkers, it is important to have a discussion between industry, academia, and regulatory agency on how to collect necessary data effectively for regulatory submission from an early stage. In general, storage of samples collected during the clinical trials is encouraged for use in retrospective PGx/biomarker analysis at a later stage if necessary. However, caution should be exercised because the sample collection process may vary widely according to the requirements of different countries. [48,49].

42.3.3 Use of PGx/Biomarkers for a Specific Drug Product and Information in the Drug Label

To use PGx/biomarker information in clinical practice, it is important how such information is included in the drug label (known as "prescribing information" in Japan). In this regard, the FDA has published a "Table of Pharmacogenomic Biomarkers in Drug Labels" (see http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm, last accessed July 18, 2013). According to recent studies [50,51], PGx/biomarker information can include genetic types of metabolizing enzymes and pharmacological targets of drugs, and the like. Examples are given in Table 42.4. The aims of using a biomarker are (1) efficacy (including information about pharmacological mechanism and/or efficacy reduction and/or lack of efficacy); and (2) safety (including information about pharmacokinetics (PK) and/or dose adjustment). The number of approved drugs with important PGx/biomarker information in the label section "Dosage/Administration" or "Indication/Usage" has increased recently [3,4,52]. Anti-cancer drugs are the most prominent ones because a target molecule of a drug (e.g., HER2, EGFR, KRAS, or ALK) is usually also described in the section "Indication/Usage" to specify the drug's target patient [53]. An increase in the number of approved drugs whose labels contain PGx/biomarker information will contribute to promoting PGx/biomarker utilization in clinical practice. Accordingly, the FDA recommends inclusion of PGx/biomarker information in drug labeling [9,24].

The importance of PGx/biomarker information on drug labels, however, can vary depending on where in the label the PGx/biomarker information has been placed (i.e., information as a reference to or information as a requirement for drug administration). Information included in what is considered a higher-level section of the label will lead to the practical use of PGx/biomarkers, but presently such cases are still limited. Thus, examining relationships between biomarkers and clinical responses (efficacy, safety, PK) during clinical trials is encouraged to provide relevant information to healthcare professionals. Furthermore, the label should be updated on a timely basis when any new scientific information on a PGx/biomarker becomes available.

The grading level of PGx/biomarker information may also differ between countries, even when the target biomarker is the same [51]. Reasons for the differences depend on how much of the collected data were from the local population, how ethnic factors influence the collected data, and what types of companion diagnostics are available. As far as the

TABLE 42.3 Examples of Biomarkers Qualified by the Regulatory Agency (EMA/FDA/PMDA)

Biomarker (Applicant)	Context of Use	Agency (Qualified Year)
7 urinary markers: • cystatin C • β2-microglobulin • total protein • clusterin • KIM-1 • TFF3 • albumin (PSTC)	• Nonclinical drug development for the detection of acute drug-induced nephrotoxicity in rats, either tubular or glomerular with associated tubular involvement. • Biomarkers provide additional and complementary information to BUN and serum creatinine to correlate with histopathological alterations.	• EMA (2009) • FDA (2008) • PMDA (2010)
Urinary clusterin (ILSI/HESI)	• May be used to detect acute drug-induced renal tubule alterations, particularly when regeneration is present, in male rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies.	• EMA (2010) • FDA (2010)
Urinary RPA-1 (ILSI/HESI)	• May be used to detect acute drug-induced renal tubular alterations, particularly in the collecting duct, in male rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies.	• EMA (2010) • FDA (2010)
CSF biomarker: • Aβ 1-42 • T-tau (BMS)	• In patients with MCI a low Aβ 1-42 and a high T-tau can help predict evolution to AD dementia type. • The CSF biomarker signature is useful for the enrichment of clinical trial populations. • To identify MCI patients who most nearly equate to the prodromal stage of AD and who are at risk for AD dementia.	EMA (2011)
Hippocampal volume (C-path CAMD)	• Low hippocampal volume, as measured by MRI and a dichotomized variable (low volume or not), appears to help enriching recruitment into clinical trials aimed at studying drugs potentially slowing the progress/conversion to AD dementia of the included subjects. • Low hippocampal volume might be considered a marker of progression to dementia in subjects with cognitive deficit compatible with predementia stage of AD, for purposes of enriching a clinical trial population.	EMA (2011)
PET amyloid imaging (BMS)	• Amyloid related positive/negative PET signal identifies patients with clinical diagnosis of predementia AD who are at increased risk for an underlying AD neuropathology, for the purpose of enriching a clinical trial population.	EMA (2012)
• CSF Aβ 1-42 • tau and/or PET amyloid imaging (BMS)	• CSF biomarker signature based on a low Aβ 1-42 and a high T-tau: to identify patients with clinical diagnosis of mild to moderate AD who are at increased risk of an underlying AD neuropathology, for purpose of enriching a clinical trial population. • Amyloid-related positive/negative PET signal to identify patients with clinical diagnosis of mild to moderate AD who are at increased risk of underlying AD neuropathology, for the purpose of enriching a clinical trial population.	EMA (2012)
• Circulating cTnT • Circulating cTnI	• Safety assessment studies in rats and dogs for monitoring and management of potential cardiotoxicity.	FDA (2012)

TABLE 42.4 Examples of PGx/Biomarkers Included in Drug Labels

Biomarker	Purpose	Characteristics	Related Drug
Metabolizing Enzyme			
CYP2D6	Safety	Drug exposure variability	Drugs metabolized by this enzyme
CYP2C9	Safety	Drug exposure variability	Drugs metabolized by this enzyme
CYP2C19	Safety	Drug exposure variability	Drugs metabolized by this enzyme
DPD	Safety	Drug exposure variability	Capecitabine, etc.
UGT1A1	Safety	Drug exposure variability	Irinotecan
Pharmacological Targets of Drugs			
HER2	Efficacy	Drug response variability	Lapatinib
EGFR	Efficacy	Drug response variability	Cetuximab
KRAS	Efficacy/Safety	Drug response variability	Panitumumab
ALK	Efficacy	Drug response variability	Crizotinib
VKORC1	Efficacy/Safety	Drug response variability	Warfarin
Others			
*HLA-B*5701*	Safety	Risk for adverse events	Abacavir
*HLA-B*1502*	Safety	Risk for adverse events	Carbamazepine

ethnic factors are concerned, differences in allele frequencies of the target PGx/biomarker among populations may affect results of clinical trials because of different PK drug profiles and/or different drug sensitivities. For example, the prevalence of the poor metabolizer allele for CYP2D6, which is well known for showing polymorphic characteristics, varies among populations [54–56]. Some other PGx/biomarkers listed in Table 42.4, such as UGT1A1, EGFR, and *HLA-B*1502*, have also been reported to show ethnic differences [35,57–59]. Therefore, to promote the proper use of a drug, the clinical impacts of PGx/biomarker information should be examined on the national population of the respective country as much as possible during drug development; otherwise, the information on the label will have very little influence in changing a physician's mind in prescribing the drug in clinical practice.

42.4 CLINICAL TRIAL DESIGN AND DRUG DEVELOPMENT UTILIZING PGx/BIOMARKER

42.4.1 Designs of Clinical Trials with PGx/Biomarkers

A PGx/biomarker can be utilized in a wide range of clinical trials, such as for patient selection (i.e., inclusion or exclusion criteria, trial enrichment or stratification), assessment of molecular mechanisms of a disease, disease state or prognosis, pharmacological action of a drug, mechanism of ADRs, and dose optimization (i.e., quantitative algorithmic determination of likely dose range), and monitoring/optimizing drug efficacy/safety (i.e., indicating or predicting ADRs, detecting their onset or reversibility) [7]. To examine the clinical utility of a PGx/biomarker in drug administration, it is very important to consider how the clinical trial should be designed. Figure 42.1 shows typical designs of clinical trials using PGx/biomarkers. Other designs are also available.

Characteristics and advantages/disadvantages of each design are described in the following subsections.

42.4.1.1 PGx/Biomarker Cohort Design

In this design, subjects are randomized independently of the results of PGx/biomarker screening (Figure 42.1, top (a)). Therefore, this study design is basically used for exploratory purposes, as in the case of identification of a safety biomarker. Any imbalance in allocating the subjects with a targeted PGx/biomarker between arms of a study may cause a bias in evaluating the PGx/biomarker data. This bias may also be caused by a limited number of collected samples (convenient sample), because only a portion of the enrolled subjects may have provided consent for the analysis [60].

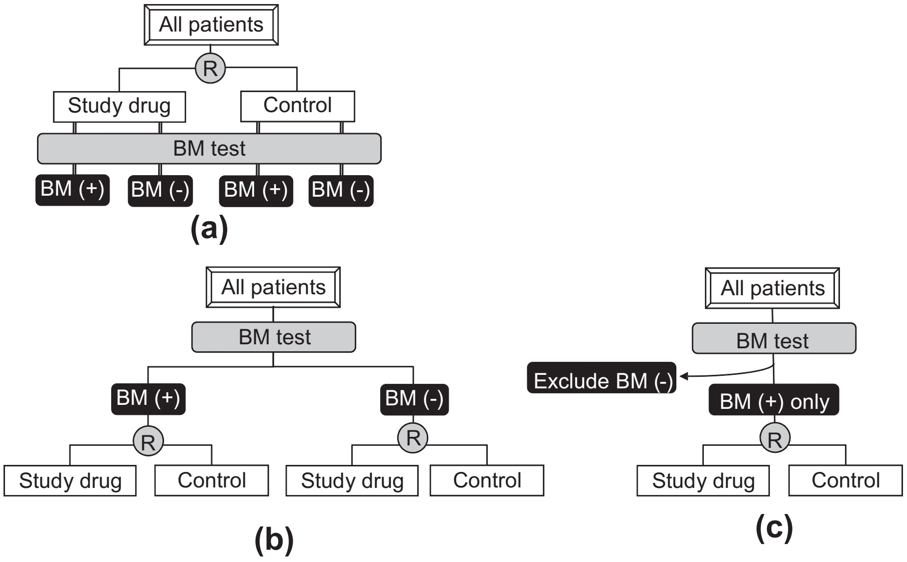


FIGURE 42.1 **Common study designs used in clinical trials utilizing PGx/biomarker.** (Top) Randomization is independent of results of PGx/BM screening. (Bottom) Randomization uses results of PGx/BM screening. R—randomization; BM (+)/(−)—biomarker positive/negative population. (a) PGx/BM cohort design; (b) PGx/BM stratified design; (c) Enriched design.

This design can be used to collect PGx/biomarker samples during a clinical trial for future analysis. It may be possible that some new scientific knowledge and/or new methodology related to PGx/biomarker, will be discovered after the clinical trial has been completed, so, it is recommended that DNA samples collected during a trial be stored properly for a certain number of years, with links to relevant clinical information, such as blood/urine drug concentration data and efficacy/safety data, so that relationships between the PGx/biomarker and drug efficacy/safety can be analyzed at a later date [12,49]. In fact, EMA, FDA, and PMDA have encouraged sample collection during clinical trials [5], and retrospective analysis-based identification of biomarkers has led to revised labeling of some drugs such as anti-cancer drugs (e.g., KRAS, EGFR) [61]. For any retrospective analysis, however, data should generally be available for all subjects (or at least for most) enrolled in a clinical trial to avoid a selection bias. A method of analysis should be predefined before conducting the retrospective analysis. Replicating the results using independent samples may help in data interpretation.

42.4.1.2 *PGx/Biomarker Stratified Design*

In this design, both biomarker test-positive and test-negative subjects are randomized based on the results of PGx/biomarker screening (Figure 42.1 bottom, (b)) [62]. This design seems to be the most rigorous in identifying the optimal target population for a drug; however, it may be difficult to apply when there are a large number of treatment arms. Sample collection for future use is also possible in this design.

42.4.1.3 *Enriched Design*

In this design, the results of PGx/biomarker screening are used for randomization as is shown in Figure 42.1, bottom, (c). The main difference is that patients without the target PGx/biomarker (PGx/biomarker (−)) are included in a stratified design but are excluded from the "enriched design" in the clinical trial. This design is typically used for a molecular-targeted drug like anti-cancer drugs [53]. Here, a prespecified cut-off value of the biomarker is used to select a subject to enroll in the study. Subjects who are likely to respond to the drug are usually selected. In some cases, subjects who have a high risk of ADR can be excluded with the help of the PGx/biomarker's cut-off value. In particular, this design is sometimes adopted when the target PGx/biomarker test negative population is expected to have a higher risk of ADR. The "enriched design," however, cannot validate the target biomarker itself, because it does not ensure that the drug is effective or safe only in PGx/biomarker test-positive subjects and the same efficacy/safety may be obtained even in PGx/biomarker test-negative subjects. Therefore, clinical impacts of excluding the target PGx/biomarker test-negative population from the trial should be considered carefully, because such an approach can deprive patients of an opportunity to receive drug therapy [3]. Sample collection for future use is also possible in this design.

As described, each design has advantages and disadvantages [3,13,63], and these should be thoroughly considered while planning a PGx/biomarker-guided trial. All data and information available at the time of the trial should be systematically collected and carefully evaluated to design appropriate clinical trials. The FDA's guidelines describe some other points to consider in designing a clinical trial with PGx/biomarkers [13,63].

42.4.2 Study Design and Development Strategy for PGx/Biomarker Analysis

42.4.2.1 PK Evaluation with PGx/Biomarkers

PGx/biomarker effects on PK have been actively studied and have been well recognized in the last decade. For example, a significant difference in PK between subjects can be explained by polymorphisms of drug-metabolizing enzymes, such as cytochrome P450 [64,65]. Thus, in general, the effects of genetic polymorphisms of the drug-metabolizing enzyme of the anticipated drug on its PK should be examined at an early stage of drug development. Although the "PGx/biomarker cohort design" can be used to examine the effects of genetic polymorphisms on PK for exploratory purposes, the PGx/biomarker stratified design is recommended when larger interindividual differences due to the drug-metabolizing enzyme's genetic polymorphism are expected.

When data from two or more studies are compared or integrated to investigate the effects of genetic factors on PK, any differences in the study designs, including inclusion/exclusion criteria, sample collection, and PGx/biomarker analysis, as well as PK measurement, should be carefully evaluated to avoid data misinterpretation. For example, it would be difficult to conclude the effects of genetic factors on PK based on data from two or more independent studies, if PK was measured using methods with different assay sensitivity or different standards for sample collection. Clinical trial site and year of a study can also affect results. More attention should be paid especially when data from different countries/regions are used for the analysis, because extrinsic factors and intrinsic factors might be significantly different between these countries/regions, which in turn may have different impacts on study results. This point should also be taken into consideration in other types of clinical trials, which are described next.

42.4.2.2 Exploratory Clinical Trials with PGx/Biomarkers

In the exploratory stage, the use of PGx/biomarkers in clinical trials is strongly encouraged. At this stage, the effects of a PGx/biomarker on drug efficacy are usually investigated because a relatively small number of patients are enrolled in the clinical trial. In addition, sometimes the effects of genetic factors on drug safety can be examined when a strong association is anticipated between a PGx/biomarker and drug safety. Results of earlier studies examining the clinical impact of a PGx/biomarker on drug efficacy/safety are very important in designing a clinical trial using this PGx/biomarker at a later stage. For example, studies on the effects of genetic factors on the dose/exposure–response relationship would provide critical information in deciding on an optimal dose for a target population. For a molecular-targeted drug such as a tyrosine kinase inhibitor in oncology, differences in efficacies between the target gene test-positive and test-negative population should be examined as early as possible during the course of drug development. Although the PGx/biomarker cohort design can be useful in establishing a possible hypothesis regarding an optimal dose or a target population, PGx/biomarker stratified design has potential to formulate more reliable hypotheses [3,13]. The enriched design can also be used in cases where a particular PGx/biomarker is expected to be clearly associated with drug efficacy or safety.

42.4.2.3 Confirmatory Clinical Trials with PGx/Biomarkers

The clinical utility of PGx/biomarkers in drug administration, based on the hypothesis established in an exploratory stage, should finally be examined in a confirmatory clinical trial. This step is crucial to the use of this PGx/biomarker in clinical practice. The PGx/biomarker stratified design can be a standard method for this purpose [62]. The enriched design can also be used if sufficient evidence was collected at the exploratory stage to enroll the target population using the PGx/biomarker, such as in the case of gefitinib administration in patients with non-small-cell lung cancer with mutated EGFR [66]. Data from a retrospective analysis may also be useful to support data interpretation.

42.4.2.4 Postapproval Clinical Trials with PGx/Biomarkers

At the time of drug approval, safety data are usually limited mainly because there were limited numbers of subjects enrolled in the clinical trials carried out before the drug was approved. Generally, an approved drug is administered to a broader population in clinical practice (after approval) than in clinical trials (before approval). Thus, rare but serious ADRs (e.g., severe skin or hepatic reactions) can be discovered and characterized after approval, when more people are taking the drug. Similarly, new PGx/biomarker-related scientific knowledge and/or a new methodology can be discovered after approval. For these reasons, the usefulness of a PGx/biomarker should be examined even in the postapproval stage. At this stage, it may be possible to find, for example, a safety risk-associated PGx/biomarker, as was reported in the case of *HLA-B*1502*, which is associated with carbamazepine-induced SJS [57], and UGT1A1, which is associated with irinotecan-induced neutropenia [67]. At this stage, retrospective analysis of stored samples can be useful in examining possible involvement of PGx/biomarkers in drug efficacy/safety. If a possibility is identified, a prospective analysis should be carried out.

42.5 FUTURE CHALLENGES

Although PGx/biomarker-guided development and administration of drugs have recently increased, there still remain many challenges with the use of PGx/biomarkers in clinical practice [3].

The first challenge is how to promote sample collection and storage for future PGx/biomarker analysis, considering that the current rate of sample collection in clinical trials is relatively low (50–69%) for most therapeutic areas [49]. A recent workshop clearly identified several ways to increase collection, including (1) international guidance on sample collection for future research, (2) additional discussion related to coding, scope, and return of results, (3) additional instruction/education for future research on PGx/biomarkers and long-term specimen storage [68]. Recent discussion on regulatory guidelines may clear the way for international harmonization [5].

The second challenge is how to increase the number of internationally qualified PGx/biomarkers. Currently, the number qualified by a single regulatory agency is limited, and the number qualified by multiple international regulatory agencies is almost none (Table 42.3). Biomarkers for nephrotoxicity in rats are probably the first ones to be qualified by all ICH regulatory agencies (EMA, FDA, and PMDA) [43]. In the era of global drug development, international qualification is absolutely necessary; it can be facilitated by submitting PGx/biomarker-related data simultaneously to the regulatory agencies of multiple countries. This may be the key to reinforcing regulatory cooperation between these international regulatory agencies with sufficient resources for establishing a novel qualification process among them.

The third challenge is about how ethnic factors influence PGx/biomarker data. There are indeed some examples where ethnic factors affect proper data evaluation. For example, the *HLA-B*1502* allele is important for predicting carbamazepine-induced SJS and TEN in Han Chinese, but not in Japanese patients [35,57,69]. Similarly, *UGT1A1*6* is clinically important when administering irinotecan as it is known to induce severe neutropenia in Japanese but not in Caucasian patients [58,70]. It is important to carefully evaluate whether the role of a biomarker in one population can be extrapolated to others with respect to ethnic factors. Examining the effects of a PGx/biomarker on drug efficacy/safety through global clinical trials is probably the best possible method of accumulating scientific evidence and evaluating how ethnic factors might influence the outcome.

The fourth challenge is how to promote co-development of companion diagnostics (CDx). Drugs that require the use of a particular diagnostic device (i.e., CDx) for ensuring balance between presumed benefits and risks have become relatively common. Although several CDxs, each specifying a target population in terms of efficacy, have been approved in oncology, such as CDxs for CCR4 and ALK [3], their numbers are still limited. To promote PGx/biomarker-guided drug administration in clinical practice, a CDx should be approved simultaneously with the drug. This concept is clearly described in the EMA, FDA, and PDMA regulatory guidelines [15,16,18]. However, synchronized development of a CDx and a drug remains a major roadblock. An early-stage close collaboration between the pharmaceutical and CDx development teams/companies may be the answer, as reported for the anti-cancer drug vemurafenib (BRAF kinase inhibitor) [71]. In the near future, establishing international CDx guidelines will be necessary to appropriately promote the co-development process at the global level.

The fifth challenge is how to provide a drug to patients who are off-target for a PGx/biomarker. The enrichment approach may be an effective way to clearly identify a drug's target population and to increase clinical trial success rates. However, a patient selection process without sufficient supporting data may lead to an inappropriate drug development strategy, in which PGx/biomarker test-negative patients would miss a chance to be treated with the drug [3]. Therefore, all available data should be carefully evaluated and the appropriateness of the strategy for enrichment should be presented to the regulatory agency for review.

In conclusion, all of the challenges just described must be overcome if PGx/biomarkers are to be properly utilized in clinical practice. As shown in Figure 42.2, they are interrelated. To better accumulate PGx/biomarker data, samples should be efficiently collected during clinical trials, as successful qualification of a biomarker depends on how much information is available. A CDx should be developed as early as possible for a qualified context of drug administration, and this context should be included in the drug label. If target patients are identified, steps should be taken to provide a drug to off-target patients. Finally, it is vital to that the challenges encountered from the early stages of drug development be discussed with a regulatory agency for the process to succeed.

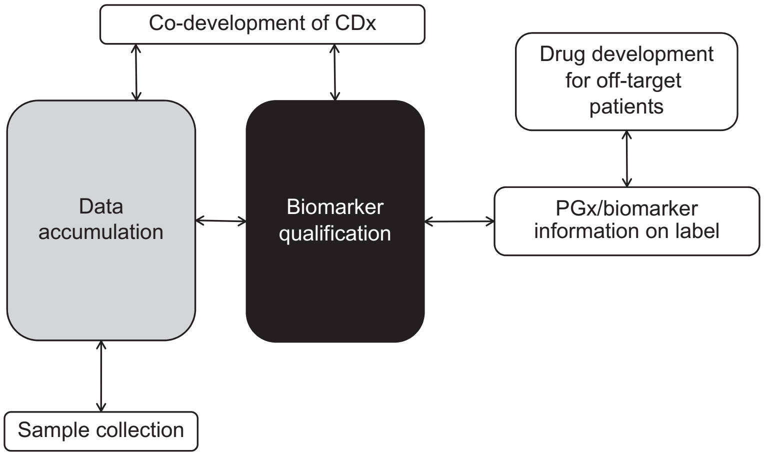


FIGURE 42.2 **Relationships among challenges for PGx/biomarker utilization in clinical practice.** Arrows indicate how each factor facilitates other factors. Every item can be not only a goal but also a means.

DISCLAIMER

The views expressed in this chapter are those of the authors and do not necessarily reflect the official views of Japanese Pharmaceuticals and Medical Devices Agency.

GLOSSARY

Biomarker a measurable characteristic that is an indicator of normal biologic processes, pathogenic processes, and/or response to therapeutic or other interventions

Companion diagnostics an *in vitro* diagnostic device that provides information essential for the safe and effective use of a corresponding therapeutic product

Label a description of a drug product, including what it is used for, who should take it, side effects, instructions for use, and safety information

Pharmacogenetics (PGt) the study of variations in DNA sequence as related to drug response

Pharmacogenomics (PGx) the study of variations in DNA and RNA characteristics as related to drug response

Qualification a conclusion that, within the stated context of use, the results of assessment with a biomarker can be relied on to adequately reflect or forecast a biological process, response, or event, and to support use of the biomarker during new drug development, ranging from discovery through postapproval

ACRONYMS AND ABBREVIATIONS

AD	Alzheimer's disease
ADME	absorption, distribution, metabolism, and excretion
ADRs	adverse drug reactions
ALK	anaplastic lymphoma kinase
Aβ	amyloid beta
BLA	Biologics License Application
BMS	Bristol-Myers Squibb
CCR4	CC chemokine receptor 4
CDx	companion diagnostics
CHMP	Committee for Medicinal Products for Human Use
C-path CAMD	Critical path coalition against major diseases
CSF	cerebrospinal fluid
cTnI	cardiac troponins I
cTnT	cardiac troponins T
CYP	cytochrome P450
DNA	deoxyribonucleic acid
DPD	dihydropyrimidine dehydrogenase
EGFR	epidermal growth factor receptor
EMA	European Medicines Agency
FDA	Food and Drug Administration
GLP	good laboratory practice
HER2	human epidermal growth factor receptor type2
HLA	human leukocyte antigen
ICH	International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use
ILSI/HESI	International Life Sciences Institute/Health and Environmental Sciences Institute
IND	investigational new drug
KIM-1	kidney injury molecule-1
KRAS	Kirsten rat sarcoma 2 viral oncogene homolog
MCI	mild cognitive impairment
MHLW	Ministry of Health, Labor and Welfare
MRI	magnetic resonance image
mRNA	messenger ribonucleic acid
NDA	new drug application
PD	pharmacodynamics
PET	positron emission tomography
PgWP	Pharmacogenomics Working Party
PGx	pharmacogenomics
PK	pharmacokinetics
PMDA	Pharmaceuticals and Medical Devices Agency
POP	PMDA Omics Project
PSTC	Predictive Safety Testing Consortium
RPA-1	Renal papillary antigen-1
SJS	Stevens-Johnson syndrome
TEN	toxic epidermal necrolysis
TFF3	Trefoil factor 3

T-tau	total tau
UGT	uridine diphosphate-glucuronosyltransferase
VKORC	vitamin K epoxide reductase complex subunit 1
VXDS	voluntary exploratory data submission

REFERENCES

[1] Hayashi K, Masuda S, Kimura H. Impact of biomarker usage on oncology drug development. J Clin Pharm Ther 2013;38(1):62–7.

[2] Walker I, Newell H. Do molecularly targeted agents in oncology have reduced attrition rates? Nat Rev Drug Discov 2009;8(1):15–16.

[3] Otsubo Y, Ishiguro A, Uyama Y. Regulatory perspective on remaining challenges for utilization of pharmacogenomics-guided drug developments. Pharmacogenomics 2013;14(2):195–203.

[4] Frueh FW, Amur S, Mummaneni P, Epstein RS, Aubert RE, DeLuca TM, et al. Pharmacogenomic biomarker information in drug labels approved by the United States food and drug administration: prevalence of related drug use. Pharmacotherapy 2008;28(8):992–8.

[5] Maliepaard M, Nofziger C, Papaluca M, Zineh I, Uyama Y, Prasad K, et al. Pharmacogenetics in the evaluation of new drugs: a multiregional regulatory perspective. Nat Rev Drug Discov 2013;12(2):103–15.

[6] ICH E15 guideline; definitions for genomic biomarkers, pharmacogenomics, pharmacogenetics, genomic data and sample coding categories; 2007. <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E15/Step4/E15_Guideline.pdf> [cited 18.07.13].

[7] ICH E16 guideline; biomarkers related to drug or biotechnology product development: context, structure and format of qualification submissions; 2010. <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E16/Step4/E16_Step_4.pdf> [cited 18.07.13].

[8] European Medicines Agency. Guideline on the investigation of drug interactions [CPMP/EWP/560/95/Rev. 1 Corr.*]; 2012. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500129606.pdf> [cited 18.07.13].

[9] Food and Drug Administration. Draft guidance for industry: drug interaction studies – study design, data analysis, implications for dosing, and labeling recommendations; 2012. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm292362.pdf> [cited 18.07.13].

[10] Ministry of Health, Labour and Welfare. Guideline on clinical pharmacokinetic studies of pharmaceuticals (796; evaluation and licensing division, pharmaceutical bureau); 2001. <http://wwwhourei.mhlw.go.jp/cgi-bin/t_docframe2.cgi?MODE=tsuchi&DMODE=SEARCH&SMODE=NORMAL&KEYWORD=%88%e3%96%f2%95%69%82%cc%97%d5%8f%b0%96%f2%95%a8%93%ae%91%d4&EFSNO=4359&FILE=FIRST&POS=0&HITSU=4> [cited 18.07.13].

[11] Ministry of Health, Labour and Welfare. Guideline on methods of drug interaction study (813, evaluation and liscensing division, pharmaceutical bureau); 2001. <http://wwwhourei.mhlw.go.jp/cgi-bin/t_docframe2.cgi?MODE=tsuchi&DMODE=SEARCH&SMODE=NORMAL&KEYWORD=%88%e3%96%f2%95%69%82%cc%97%d5%8f%b0%96%f2%95%a8%93%ae%91%d4&EFSNO=4360&FILE=FIRST&POS=0&HITSU=4> [cited 18.07.13].

[12] European Medicines Agency. Reflection paper on pharmacogenomic samples, testing and data handling (EMEA/CHMP/PGxWP/201914/2006); 2007. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003864.pdf> [cited 18.07.13].

[13] Food and Drug Administration. Guidance for industry: clinical pharmacogenomics: premarket evaluation in early-phase clinical studies and recommendations; 2013. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM337169.pdf> [cited 18.07.13].

[14] Ministry of Health, Labour and Welfare. General principles for clinical trials using pharmacogenomics (0930007, evaluation and licensing division, pharmaceuticals and food safety beureau); 2008. <http://www.pmda.go.jp/operations/shonin/outline/shinrai/file/tuchi/1_0930007.pdf> [cited 18.07.13].

[15] Ministry of Health, Labour and Welfare. General points to consider in the regulatory submission of a drug and companion diagnostics for approval (0712-12, evaluation and licensing division, pharmaceuticals and food safety beureau); 2013. <http://wwwhourei.mhlw.go.jp/hourei/doc/tsuchi/T130702I0070.pdf> [cited 18.07.13].

[16] European Medicines Agency. Reflection paper on co-development of pharmacogenomic biomarkers and assays in the context of drug development; draft (EMA/CHMP/641298/2008); 2010. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/07/WC500094445.pdf> [cited 18.07.13].

[17] Food and Drug Administration. Guidance for industry and FDA staff: pharmacogenetic tests and genetic tests for heritable markers; 2007. <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM071075.pdf> [cited 18.07.13].

[18] Food and Drug Administration. Draft guidance for industry and food and drug administration staff: *in vitro* companion diagnostic devices; 2011. <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM262327.pdf> [cited 18.07.13].

[19] Food and Drug Administration and European Medicines Agency. General principles: processing joint FDA EMEA Voluntary Genomic Data Submissions (VGDSs) within the framework of the confidentiality arrangement; 2006. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003887.pdf> [cited 18.07.13].

[20] Food and Drug Administration. Guidance for industry: pharmacogenomic data submissions; 2005. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM079849.pdf> [cited 18.07.13].

[21] Food and Drug Administration. Draft guidance for industry: pharmacogenomic data submissions—companion guidance; 2007. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM079855.pdf> [cited 18.07.13].

[22] European Medicines Agency. Reflection paper on methodological issues associated with pharmacogenomic biomarkers in relation to clinical development and patient selection; draft (EMA/446337/2011); 2011. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/07/WC500108672.pdf> [cited 18.07.13].

[23] European Medicines Agency. Guideline on the use of pharmacogenetic methodologies in the pharmacokinetic evaluation of medicinal products (EMA/CHMP/37646/2009); 2012. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/02/WC500121954.pdf> [cited 18.07.13].

[24] Food and Drug Administration. Draft guidance for industry: clinical pharmacology section of labeling for human prescription drug and biological products – content and format; 2009. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm109739.pdf> [cited 18.07.13].

[25] European Medicines Agency. Reflection paper on pharmacogenomics in oncology; draft (EMEA/CHMP/PGxWP/128435/2006); 2008. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003866.pdf> [cited 18.07.13].

[26] European Medicines Agency. Reflection paper on the use of genomics in cardiovascular clinical intervention trials (EMEA/CHMP/PGxWP/278789/2006); 2007. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003865.pdf> [cited 18.07.13].

[27] Ministry of Health, Labour and Welfare. Guideline on clinical evaluation of anti-cancer drugs; 2005. <http://www.pref.chiba.lg.jp/yakumu/iyakubugaihin/documents/aku.pdf> [cited 18.07.13].

[28] European Medicines Agency. Concept paper on key aspects for the use of pharmacogenomic methodologies in the pharmacovigilance evaluation of medicinal products (EMA/CHMP/917570/2011); 2011. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/01/WC500120775.pdf> [cited 18.07.13].

[29] Ministry of Health, Labour and Welfare. Request to cooperate in research regarding severe adverse reactions (skin disorder, rhabdomyolysis and interstitial pneumonia) (0926-2, safety division, pharmaceuticals and food safety bureau); 2011. <http://www.pref.ibaraki.jp/bukyoku/hoken/itakohc/topic/201110/yuugai.pdf> [cited 18.07.13].

[30] European Medicines Agency. Mandate, objectives and rules of procedure for the CHMP PG working party (EMEA/CHMP/125959/2008 Rev.1); 2009. <http://www.ema.europa.eu/docs/en_GB/document_library/Other/2010/01/WC500069715.pdf> [cited 18.07.13].

[31] European Medicines Agency. Report to the CPMP on the EMEA seminar on the use of pharmacogenetics in the drug development process (EMEA/CPMP/1483/00); 2000. <www.ema.europa.eu/docs/en_GB/document_library/Report/2010/01/WC500069722.pdf> [cited 18.07.13].

[32] Ministry of Health, Labour and Welfare. Submission of information to regulatory authority for preparation of guidance on the use of pharmacogenomics in clinical trials (0318001, evaluation and licensing division, pharmaceuticals and food safety beureau); 2005. <http://www.jhsf.or.jp/hsmag/publish/info_20050318_01.pdf> [cited 18.07.13].

[33] Ministry of Health, Labour and Welfare. Basic principles on global clinical trials (0928010, evaluation and licensing division, pharmaceuticals and food safety bureau); 2007. <http://www.pmda.go.jp/kijunsakusei/file/guideline/new_drug/GlobalClinicalTrials_en.pdf> [cited 18.07.13].

[34] Ministry of Health, Labour and Welfare. Basic principles on global clinical trials (reference cases) (administrative notice, evaluation and licencing division, pharmaceuticals and food safety bureau); 2012. <http://www.pmda.go.jp/kijunsakusei/file/guideline/new_drug/GCT-jirei_en.pdf> [cited 18.07.13].

[35] Kaniwa N, Saito Y, Aihara M, Matsunaga K, Tohkin M, Kurose K, et al. HLA-B locus in Japanese patients with anti-epileptics and allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis. Pharmacogenomics 2008;9(11):1617–22.

[36] Kaniwa N, Saito Y, Aihara M, Matsunaga K, Tohkin M, Kurose K, et al. HLA-B*1511 is a risk factor for carbamazepine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in Japanese patients. Epilepsia 2010;51(12):2461–5.

[37] Tohkin M, Kaniwa N, Saito Y, Sugiyama E, Kurose K, Nishikawa J, et al. A whole-genome association study of major determinants for allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis in Japanese patients. Pharmacogenomics J 2013;13(1):60–9.

[38] Food and Drug Administration. Innovation or stagnation: challenge and opportunity on the critical path to new medical products; 2004. <http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.html> [cited 18.07.13].

[39] Thiers FA, Sinskey AJ, Berndt ER. Trends in the globalization of clinical trials. Nat Rev Drug Discov 2008;7(1):13–14.

[40] Khin NA, Yang P, Hung HMJ, Maung-U K, Chen YF, Meeker-O'Connell A, et al. Regulatory and scientific issues regarding use of foreign data in support of new drug applications in the United States: an FDA perspective. Clin Pharmacol Ther 2013.

[41] Ando Y, Uyama Y. Multiregional clinical trials: Japanese perspective on drug development strategy and sample size for Japanese subjects. J Biopharm Stat 2012;22(5):977–87.

[42] Goodsaid F. Challenges of biomarkers in drug discovery and development. Expert Opin Drug Discov 2012;7(6):457–61.

[43] Dieterle F, Sistare F, Goodsaid F, Papaluca M, Ozer JS, Webb CP, et al. Renal biomarker qualification submission: a dialog between the FDA-EMEA and predictive safety testing consortium. Nat Biotechnol 2010;28(5):455–62.

[44] Pharmaceuticals and Medical Devices Agency. Record of the consultation on pharmacogenomics/biomarkers for drug-induced kidney injury biomarkers; 2010. <http://www.pmda.go.jp/english/service/biomarker.html> [cited 01.08.13].

[45] Isaac M, Vamvakas S, Abadie E, Jonsson B, Gispen C, Pani L. Qualification opinion of novel methodologies in the predementia stage of Alzheimer's disease: Cerebro-spinal-fluid related biomarkers for drugs affecting amyloid burden – regulatory considerations by European Medicines Agency focusing in improving benefit/risk in regulatory trials. Eur Neuropsychopharmacol 2011;21(11):781–8.

[46] European Medicines Agency. Qualification opinion of low hippocampal volume (atrophy) by MRI for use in clinical trials for regulatory purpose – in pre-dementia stage of Alzheimer's disease (EMA/CHMP/SAWP/809208/2011); 2011. <http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2011/12/WC500118737.pdf> [cited 22.07.13].

[47] European Medicines Agency. Qualification opinion of Alzheimer's disease novel methodologies/biomarkers for PET amyloid imaging (positive/negative) as a biomarker for enrichment, for use in regulatory clinical trials in predementia Alzheimer's disease (EMA/CHMP/SAWP/892998/2011); 2011. <http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2012/04/WC500125018.pdf> [cited 30.07.13].

[48] Ricci DS, Broderick ED, Tchelet A, Hong F, Mayevsky S, Mohr DM, et al. Global requirements for DNA sample collections: results of a survey of 204 ethics committees in 40 countries. Clin Pharmacol Ther 2011;89(4):554–61.

[49] Warner AW, Bhathena A, Gilardi S, Mohr D, Leong D, Bienfait KL, et al. Challenges in obtaining adequate genetic sample sets in clinical trials: the perspective of the industry pharmacogenomics working group. Clin Pharmacol Ther 2011;89(4):529–36.

[50] Kadowaki K, Ishiguro A, Takamatsu S, Saito Y, Uyama Y. Analytical consideration on pharmacogenomics information in the package inserts of drugs and availability of genetic tests in Japan [in Japanese]. Regul Sci Med Prod 2012;2(2):83–91.

[51] Otsubo Y, Asahina Y, Noguchi A, Sato Y, Ando Y, Uyama Y. Similarities and differences between US and Japan as to pharmacogenomic biomarker information in drug labels. Drug Metab Pharmacokinet 2012;27(1):142–9.

[52] Uyama Y, Ishiguro A, Nakamura H, Toyoshima S. Use of biomarker in drug development-Japanese perspectives. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2012.

[53] Ishiguro A, Yagi S, Uyama Y. Characteristics of pharmacogenomics/biomarker-guided clinical trials for regulatory approval of anti-cancer drugs in Japan. J Hum Genet 2013;58(6):313–6.

[54] Kurose K, Sugiyama E, Saito Y. Population differences in major functional polymorphisms of pharmacokinetics/pharmacodynamics-related genes in Eastern Asians and Europeans: implications in the clinical trials for novel drug development. Drug Metab Pharmacokinet 2012;27(1):9–54.

[55] Zhou S-F. Polymorphism of human cytochrome P450 2D6 and its clinical significance: Part I. Clin Pharmacokinet 2009;48(11):689–723.

[56] Zhou S-F. Polymorphism of human cytochrome P450 2D6 and its clinical significance: Part II. Clin Pharmacokinet 2009;48(12):761–804.

[57] Chung WH, Hung SI, Hong HS, Hsih MS, Yang LC, Ho HC, et al. Medical genetics: a marker for Stevens-Johnson syndrome. Nature 2004;428(6982):486.

[58] Kaniwa N, Kurose K, Jinno H, Tanaka-Kagawa T, Saito Y, Saeki M, et al. Racial variability in haplotype frequencies of UGT1A1 and glucuronidation activity of a novel single nucleotide polymorphism 686C>T (P229L) found in an African-American. Drug Metab Dispos 2005;33(3):458–65.

[59] Saijo N, Fukuoka M, Thongprasert S, Ichinose Y, Mitsudomi T, Mok TSK, et al. Lung cancer working group report. Jpn J Clin Oncol 2010;40(Suppl. 1):i7–i12.

[60] Wang S-J, O'Neill RT, Hung HJ. Statistical considerations in evaluating pharmacogenomics-based clinical effect for confirmatory trials. Clin Trials 2010;7(5):525–36.

[61] Becquemont L, Alfirevic A, Amstutz U, Brauch H, Jacqz-Aigrain E, Laurent-Puig P, et al. Practical recommendations for pharmacogenomics-based prescription: 2010 ESF-UB conference on pharmacogenetics and pharmacogenomics. Pharmacogenomics 2011;12(1):113–24.

[62] Ziegler A, Koch A, Krockenberger K, Großhennig A. Personalized medicine using DNA biomarkers: a review. Hum Genet 2012;131(10):1627–38.

[63] Food and Drug Administration. Draft guidance for industry: enrichment strategies for clinical trials to support approval of human drugs and biological products; 2012. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf> [cited 18.07.13].

[64] Kirchheiner J, Fuhr U, Brockmoller J. Pharmacogenetics-based therapeutic recommendations–ready for clinical practice? Nat Rev Drug Discov 2005;4(8):639–47.

[65] Tomalik-Scharte D, Lazar A, Fuhr U, Kirchheiner J. The clinical role of genetic polymorphisms in drug-metabolizing enzymes. Pharmacogenomics J 2007;8(1):4–15.

[66] Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H, et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med 2010;362(25):2380–8.

[67] Minami H, Sai K, Saeki M, Saito Y, Ozawa S, Suzuki K, et al. Irinotecan pharmacokinetics/pharmacodynamics and UGT1A genetic polymorphisms in Japanese: roles of UGT1A1*6 and *28. Pharmacogenet Genomics 2007;17:497–504.

[68] Warner AW, Bienfait KL, Bledsoe M, Burckart G, Flamion B, Knoppers B, et al. Improving clinical trial sampling for future research—an international approach: outcomes and next steps from the DIA future use sampling workshop 2011. Pharmacogenomics 2013;14(1):103–12.

[69] Chen P, Lin JJ, Lu CS, Ong CT, Hsieh PF, Yang CC, et al. Carbamazepine-induced toxic effects and HLA-B*1502 screening in Taiwan. N Engl J Med 2011;364(12):1126–33.

[70] Onoue M, Terada T, Kobayashi M, Katsura T, Matsumoto S, Yanagihara K, et al. UGT1A1*6 polymorphism is most predictive of severe neutropenia induced by irinotecan in Japanese cancer patients. Int J Clin Oncol 2009;14(2):136–42.

[71] Cheng S, Koch WH, Wu L. Co-development of a companion diagnostic for targeted cancer therapy. New Biotechnology 2012;29(6):682–8.

[72] Ministry of Health, Labour and Welfare. Technical guidance on development of *in vitro* companion diagnostics and corresponding therapeutic products; 2013. <http://www.pmda.go.jp/kijunsakusei/file/companion/companion20131226.pdf> [cited 17.02.14].

This page is intentionally left blank

Part IX

Global Perspective

This page is intentionally left blank

Chapter 43

Population Diversity and Pharmacogenomics in Africa

Collen Masimirembwa,[a] Collet Dandara,[b] and Julia Hasler[a]

[a]*African Institute of Biomedical Science and Technology, Zimbabwe,* [b]*Department of Clinical Laboratory Science, University of Cape Town, South Africa*

Chapter Outline

43.1 Overview 971
- 43.1.1 Population Diversity and Epigenetics 972

43.2 Population Diversity in Africa 973
- 43.2.1 Historical Approaches 973
- 43.2.2 Africa's Rich Cultural and Natural History 973
- 43.2.3 Population Diversity and Stratified Medicine 978

43.3 Population Genetics, Diversity, Ancestry, and Admixture 979
- 43.3.1 Genetic Variation and Diversity in African Populations 979
- 43.3.2 High Genetic Diversity in African Populations 979

43.4 Pharmacogenomics in Africa 985
- 43.4.1 Pharmacogenetic Diversity in African Populations 985

43.5 Future Perspectives 992
- 43.5.1 Epigenetics 992
- 43.5.2 From Pharmacogenetics to Pharmacogenomics: Where Does Africa Feature? 993
- 43.5.3 Catalyzing Genomics Research in Africa: Human Heredity and Health 994

43.6 Conclusions 994

Key Concepts 994

Acknowledgments 995

Glossary 995

Acronyms and Abbreviations 995

References 996

43.1 OVERVIEW

This chapter provides insight into the great diversity of African populations and the implications for pharmacogenomics and stratified medicine, both for people living in Africa and for the African diaspora. Such diversity manifests itself through genetic, ethnolinguistic, and cultural differences as well as differences that arise from environmental and sociopolitical factors such as exposure to pathogens, disease burden, diet, climate, poverty, urban versus rural lifestyles, and so on.

Overwhelming archaeological and other evidence supports the thesis that Homo sapiens developed in Central-Eastern Africa about 200,000 years ago and that various migrations occurred from Africa to Europe, Asia, and the Americas 40,000–100,000 years ago [1–3]. As humans adapted to various diets, diseases, and climates, inherited traits emerged giving rise to distinct population groups with physical and physiological differences, including response to xenobiotic challenge. The greatest diversity has remained in Africa, with emigrated populations showing less diversity in proportion to the distance of migration [4,5].

Within Africa itself, various migrations took place resulting in the expansion of Bantu speakers from West-Central Africa (now Nigeria and Cameroon)—first, toward the equatorial rainforests and then into eastern and southern Africa; additionally there was migration both westward and eastward from the area that is now Sudan and migration northward of Cushitic speakers from the Ethiopian highlands to the Red Sea coast of Sudan as well as southward into Kenya and Tanzania [20] (Figure 43.1).

Early studies on the development of human diversity in Africa and elsewhere relied primarily on archaeological and linguistic evidence. With the advent of DNA sequencing, data on ethnic differences in single genes (e.g., for the lactase gene) allowed the construction of evolutionary trees

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00043-8
© 2014 Elsevier Inc. All rights reserved.

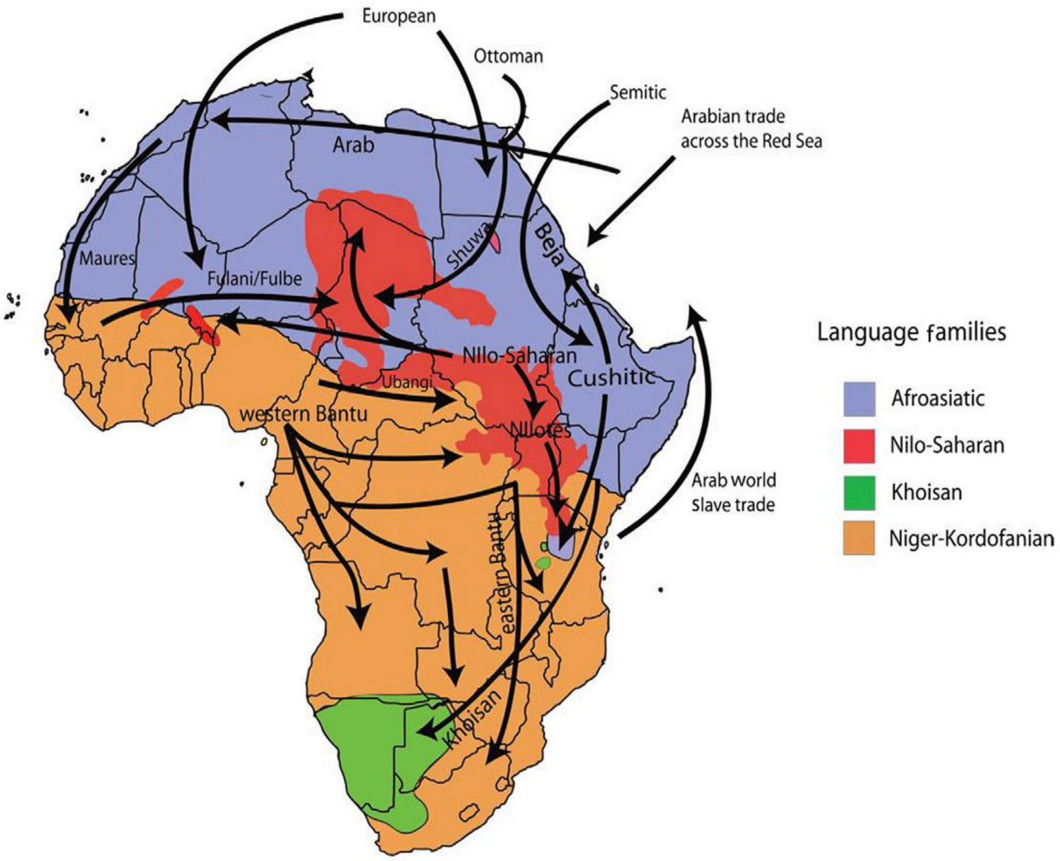


FIGURE 43.1 **Major population groups in Africa.** Data are according to ethnolinguistic groups and postulated routes of major population migrations within the African continent and later reverse migrations of non-Africans back to the continent. *Source: Reprinted with permission from Elsevier: Campbell and Tishkoff* [2].

based on families of genes. The rapidity of whole-genome sequencing has allowed genomic comparison of different ethnic groups and characterization of the genetic diversity between populations. To complement this growing body of knowledge from the biological sciences, research findings of historians, linguists, economists, geographers, sociologists, and anthropologists are required to improve our understanding of human diversity and its true impact on health and, more specifically, on pharmacogenomics and stratified medicine.

43.1.1 Population Diversity and Epigenetics

The question arises as to why one might consider broad population diversity (as opposed to genetic diversity only) in relation to pharmacogenomics and stratified medicine. As human diversity is manifested in language, cultures, habitats, and so forth, and as epigenetic modifications can result from lived experiences and environmental exposure, a multidisciplinary perspective—historical, social, cultural, linguistic, political, and economic—must be assumed.

The long-standing "nature versus nurture" debate has been challenged by the recent emergence and postulated significance of "epigenetics" [6]. The effects of lived experiences (e.g., stress, famine) in one generation upon subsequent generations have been studied in great detail by behavioral and social scientists in the context of how "nurture" affects human health and behavior. It is only in the last 10 years or so that biological scientists have been able to demonstrate that such life experiences can result in inheritable epigenetic modifications of genomic DNA, thereby providing, at least in part, an explanation for the transmission of traits associated with "nurture" [7–9].

Environmental exposures have been shown to cause epigenetic change. For example, Nieminen and colleagues [10] found distinct genetic and epigenetic signatures of colorectal cancers according to ethnic origin (i.e., Egyptian and Finnish) and concluded that the possible effects of environmental exposures in carcinogenesis occur through DNA methylation. This observation should have applications in prevention, molecular diagnosis, prognosis, and treatment. The mycotoxin fumonisin B1, found in the food chain in certain countries, displays epigenetic properties, as it has been shown to induce DNA hypermethylation [11]. On this, Lock states:

> Epigeneticists now add their considerable weight to this claim that biology must be understood as situated to the same extent as are sociohistorical forces. They have thrown down the gauntlet to genetic reductionists but, for the moment, "nurture" continues to be a poor cousin—not surprising perhaps, given the challenge of demonstrating robust evidence for molecular nature/nurture links that move us beyond mere gene–environment interactions [6].

It is therefore in this context that we structure this chapter on Africa so as to demonstrate population diversity in the widest sense, in addition to including the important area of genetic diversity. The status of pharmacogenomics in Africa is summarized; this is followed by a discussion of future challenges and priorities for research.

43.2 POPULATION DIVERSITY IN AFRICA

This section attempts to provide a broad overview of the population diversity of Africa, excluding genetic diversity. In the discussion of populations, it is important to define some terms that might have different meanings when used in different contexts. In this chapter, we address race and ethnicity as being related to biological and sociological factors, respectively. Race refers to a person's physical appearance, such as skin, eye and hair color, bone/jaw structure, and so on. Thus, the major races are Caucasian, African, and Asian. Ethnicity, on the other hand, relates to sociocultural factors such as nationality, culture, ancestry, language, and beliefs. As examples, Caucasian ethnic groups can be Irish or German; Asians ethnic groups can be Chinese or Japanese; and African ethnic groups can be Bantu, Nilotics, and so forth. The depth of ethnic classification can go further into subpopulations within a country (e.g., the Chinese can be considered an ethnic group within which are subethnicities such as Han, Zhuang, and Hui). In Africa, the use of country names to denote ethnicities faces a great challenge in that the country boundaries denote political constructs of the colonial era rather than shared characteristics typical of ethnic groups.

Race and ethnicity are both associated with genealogical links in that races claim it from geographical isolation; ethnicities, from presumed shared ancestry. With population isolation barriers gone and intermarriages common, population admixture is challenging the traditional concepts of race and ethnicity. There are, however, still identifiable frameworks of race and ethnicity that are useful in understanding human population diversity, as discussed in this chapter. Race and ethnicity, however, have also been used in political and economic conflicts so that they can evoke strong reactions in people even when used in academic endeavors such as this one. We therefore replace both with the general term *population*; that is, for the African, Asian, and Caucasian races, we refer to African, Asian, and Caucasian populations. At the ethnicity level, we also refer to populations, such as Han, Hausa, and Ibo.

43.2.1 Historical Approaches

In the context of diversity, it is important to include a note on historical approaches to the study of disease and methods for its treatment and prevention. Africa has been subject to a wide range of historical events that have had an impact on disease burden both in times past and present, some of which have resulted in genetic changes both in the host and in the parasite. In his paper on historical epidemiology and infectious disease processes in Africa, Webb states:

> *The biomedical understandings of disease processes in Africa are frequently naïve, because they lack political, social, cultural, and economic historical contexts. Africanist historians can develop these historical contexts in ways that will be relevant to the practice of public health in Africa … The importance of this work in historical epidemiology extends beyond the academy. It can raise new issues to be incorporated into disease interventions. It can inform contemporary public health planners about the consequences of past interventions* [12].

This statement surely also apply to pharmacogenomics and stratified medicine.

43.2.2 Africa's Rich Cultural and Natural History

The rich cultural and natural history of Africa is presented in a UNESCO publication [13] that highlights evidence of the remains of past civilizations and their mastery of difficult terrains, as well as of the diversity of natural habitats, wildlife, and resources from desert to tropical forest to savannah. This huge variation in natural geographic regions implies that lifestyles, diet, disease burden, and so on, varies widely across the continent depending on the region inhabited. Such environmental differences are proposed to contribute to the development of ethnic diversity [14].

43.2.2.1 Linguistic Diversity

Africa has more than 2000 different native languages, with some sources claiming nearer to 3000 [15]. The languages are grouped into several major families with a continental distribution, as shown in Figure 43.1:

- Afroasiatic (Hamito-Semitic) is spoken throughout the Middle East, North Africa, the Horn of Africa, and parts of the Sahel.
- Nilo-Saharan is spoken in Sudan and Chad.
- Niger-Congo (Niger-Kordofanian) is spoken in West, Central, and Southeast Africa.
- Khoe (Khoisan) is spoken in the deserts of Namibia and Botswana.

- Austronesian is spoken in Madagascar.
- Indo-European is spoken on the southern tip of the continent and Mauritius.

The high linguistic diversity of many African countries is best illustrated by Nigeria, which has more than 500 languages [15], and one of the greatest concentrations of linguistic diversity in the world. The Niger-Congo language family, to which Nigeria belongs, has more than 1400 languages, making it the biggest language group in the world. Some of the languages in Africa with the greatest number of speakers belong to the Niger-Congo group: these include many Bantu languages (of which the best known are Ganda, Gikuya, Kongo, Lingala, Luba-Kasai, Luyia, Mbundu (Luanda), Northern Sotho, Southern Sotho, Nyanja, Rundi, Rwanda, Shona, Sindebele, Sukuma, Swahili, Tsonga, Tswana, Umbundu, Xhosa, and Zulu; Sango, of the Central Africa Republic; Akan, the largest language group in Ghana; Yoruba and Igbo, the major languages of Nigeria; Wolof, the most common language of Senegal; Fulfulde, spoken in much of West and Central Africa; and Manding, varieties of which are found in several West African countries under various names, including Bambara, the national language of Mali; and Dyula, a widespread trade language [15].

43.2.2.2 Measuring Population Diversity

One measure for determining population diversity within a particular country or region is "ethnicity." Various measures of ethnicity have been developed since the early 1960s, with the original measure based primarily on language groups. This is the atlas *Narodov Mira*, which was developed by a Soviet team of ethnographers and primarily based on language groups. The definition of ethnicity used in developing existing measurement tools is complicated and has varied among researchers. Political scientists, social scientists, and economists have used these measures to investigate the effects of ethnic diversity on political conflict, economic growth, and social outcomes.

Fearon [16] articulated the challenges of defining ethnicity, using different approaches and a variety of sources to develop his grouping system. According to Fearon, the attribution of ethnicity is ideally carried out by a country's inhabitants. Posner [17] provides a comparative description of seven different measures, including Fearon's, and developed a system of defining ethnicity for political science research based on politically relevant ethnic groupings. A number of researchers have developed multiple measures for ethnic diversity based on different criteria for membership grouping [18,19]. Clearly, it is important for a researcher to choose a measure of ethnicity that aligns with the particular hypothesis being tested [17]. Tables 43.1 and 43.2 demonstrate the ethnic diversity of Africa based on Fearon's 2003 data [16].

Fearon's method of grouping was used in his measurement of indices for ethnic and cultural diversity. The most commonly used measure of aggregate ethnic diversity is ethnic fractionalization, defined as the probability that two individuals selected at random from a country will be from different ethnic groups. If the groups in one country speak

TABLE 43.1 Descriptive Statistics on Ethnic Groups Larger than 1% of Country Population by Region[a]

Region	Countries (n)	Ethnic Groups (n)	Ethnic Groups (% of World)	CF[b]	EF[b]
World	160	822		0.31	0.48
Western Europe/ North America/ Australia/New Zealand/Japan	21	68	13	0.19	0.24
Eastern Europe/ Former Soviet Union	31	141	19	0.30	0.41
Latin America/ Caribbean	23	84	14	0.19	0.41
Asia	23	108	14	0.33	0.44
North Africa/ Middle East	19	70	12	0.29	0.45
Sub-Saharan Africa	43	351	27	0.43	0.71

[a] *Countries with more than half a million in population in 1990.* [b] *Average fractionalization for each region: cultural (CF) and ethnic (EF).*
Source: Modified from Fearon. [16]

TABLE 43.2 Ethnic and Cultural Fractionalization in African Countries[a]

EF Ranking[b]	Country	EF	CF[b]	CF Ranking	Language Family[c]
1	Tanzania	0.953	0.564	14	NK
2	Democratic Republic of Congo	0.933	0.628	7	NK
3	Uganda	0.93	0.647	5	NS and NK
4	Liberia	0.899	0.644	6	NK
5	Cameroon	0.887	0.733	1	NK
6	Togo	0.883	0.602	8	NK
7	South Africa	0.88	0.53	20	NK and Kh
8	Congo	0.878	0.562	15	NK
9	Madagascar	0.861	0.192	38	Austronesian
10	Gabon	0.857	0.382	29	NK
11	Kenya	0.852	0.601	9	AA, NK & NS
12	Ghana	0.846	0.388	28	NK
13	Malawi	0.829	0.294	32	NK
14	Guinea Bissau	0.818	0.568	13	NK
15	Somalia	0.812	0.29	33	AA
16	Nigeria	0.805	0.66	4	NK; AA+NS in north
17	Central African Republic	0.791	0.511	21	NK; NS+AA in north
18	Ivory Coast	0.784	0.557	17	NK
19	Chad	0.772	0.727	2	NS+AA
20	Mozambique	0.765	0.285	34	NK
21	Gambia	0.764	0.548	18	NK
22	Sierra Leone	0.764	0.534	19	NK
23	Ethiopia	0.76	0.562	16	AA; some NS in southwest
24	Angola	0.756	0.242	36	NK
25	Mali	0.754	0.59	11	NK in south; AA in north; some NS
26	Senegal	0.727	0.402	25	NK
27	Zambia	0.726	0.189	39	NK; some KH in south-west
28	Namibia	0.724	0.589	12	Kh; some NK in north-west
29	Sudan and South Sudan[d]	0.708	0.698	3	AA in north; NS in south and part north; some NK in south-west NS in south and part north; some NK in south-west

(*Continued*)

TABLE 43.2 Continued

EF Ranking[b]	Country	EF	CF[b]	CF Ranking	Language Family[c]
30	Burkina Faso	0.704	0.354	31	NK; some AA in north
31	Guinea	0.669	0.49	22	NK
32	Eritrea	0.647	0.398	27	AA
33	Niger	0.637	0.6	10	AA; NS in east; NS and NK in southwest
34	Mauritius	0.632	0.448	23	Indo-European
35	Mauritania	0.625	0.272	35	AA
36	Benin	0.622	0.4	26	NK; some NS in north
37	Djibouti	0.606	0.404	24	AA
38	Morocco	0.479	0.36	30	AA
39	Zimbabwe	0.366	0.141	42	NK and Kh
40	Botswana	0.351	0.161	40	NK and Kh
41	Burundi	0.328	0.04	45	NK
42	Algeria	0.32	0.237	37	AA
43	Swaziland	0.28	0.143	41	NK
44	Lesotho	0.255	0.057	44	NK
45	Rwanda	0.18	0	47	NK
46	Egypt	0.164	0	48	AA
47	Libya	0.151	0.127	43	AA; NS in south
48	Tunisia	0.039	0.033	46	AA

[a] *Data combining Sub-Saharan Africa and North Africa; for ethnic groups larger than 1% of country population.* [b] *Fractionalization for each country: cultural (CF) and ethnic (EF).* [c] *NK—Nilo-Kordofanian; NS—Nilo-Saharan; AA—Afroasiatic; Kh—Khoisan.* [d] *Data not available for Sudan and South Sudan separately.*

Source: Modified from Fearon [16].

structurally unrelated languages, their cultural fractionalization (CF) index will be the same as their ethnic fractionalization (EF). The more similar the languages spoken by the different ethnic groups, the more the related cultural measure will be reduced below the value of EF for that country [16].

Using a global database, Fearon found 822 ethnic groups worldwide, 315 of them originating in Africa. Not only does Africa demonstrate more diversity than other regions of the world, but the extent of its diversity varies dramatically from country to country (see Table 43.2). An important point to note in considering diversity per country is that most African country boundaries are political constructs and many ethnic groups overlap national borders.

The ethnolinguistic data clearly demonstrate that greater ethnic diversity exists in Africa compared to the rest of the world and is consistent with the greater diversity of African populations as demonstrated in genetic studies [20,21]. These data have relevance for past and future studies of pharmacogenetics and pharmacogenomics, as there is a requirement to define and align ethnicity with genetic data.

43.2.2.3 *Diversity in Demographic, Social, Economic, and Health-Related Indicators*

Africa as a continent displays considerable diversity in many different indicators; this is evident from statistical data available from the World Health Organisation (WHO), the UN Department of Economic and Social Affairs (UNESA), the World Bank, the African Development Bank, and other sources. It is apparent that for the African continent "no one size fits all." This will be of importance to the stratified medicine of the future. In any research project or health intervention in an African country, the social, economic, and cultural determinants of health each must be considered. We have chosen to illustrate African diversity using the socioeconomic and health indicators in the following subsections.

Urban versus Rural

Data from the UNESA show the urban percentage of populations from all regions of the world for 2011 with estimates for 2030 [22]. As a region, Africa exhibits the greatest diversity, with an urban percentage of the populations ranging from less than 20% to greater than 80%. This clearly has implications for diet, environmental pollutant exposure, life style, and access to education and health services.

Environmental Health

Several measures are classified among environmental heath indicators (e.g., alcohol and tobacco consumption, access to sanitation and clean water, indoor and outdoor pollution). These indicators show great differences between Africa and other regions as well as differences between African countries. For example, WHO data [23] show that African countries have a marked variation in the percentage of their populations with access to clean water (30–99%) and that the continent differs markedly from most of the rest of the world in this respect. From a health perspective, lack of access to clean water increases the population risk of waterborne diseases.

Life Expectancy and Age Distribution

Life expectancy has an impact on the proportions of different age groups in a population. Africa shows the greatest diversity in this regard, with life expectancy in some countries at 47–55 years of age and others greater than 70 years. Other regions in the world are more uniform in terms of life expectancy, which averages around 70–80 [22,23]. In planning for stratified medicine in the future, different regions of the world will have to take into account the age proportions in their populations. Burden of disease and ensuing healthcare needs vary considerably depending on the profile of the different age groups. In the future, while most of the world will have to cater to a large extent to the health needs of those over 60 years of age, African countries will likely be more diverse in this regard [24].

Burden of Disease

Researchers, governments, and the pharmaceutical industry must take note of the burden of disease and how it changes over time when planning future research or health policy interventions. For future pharmacogenomics studies and for progress toward stratified medicine, decisions have to be made with respect to prioritizing diseases and standardizing treatment regimens. In this context, it is useful to note the most recently available comprehensive health-related statistics compiled by WHO by region and by country [25,26]. Examples of the WHO data are shown in Tables 43.3–43.5.

What is interesting about the WHO data in Table 43.3 is that Africa has the highest mortality rates (30–70 age group) for cardiovascular disease and diabetes, and that the mortality rates for communicable and noncommunicable diseases (NCDs) are more or less equal. Cancer, cardiovascular disease, and diabetes have been considered diseases of the developed world, but figures show that attention to them is also required in Africa. The implications of these diseases are that the drugs for treatment of NCDs in Africans should not be neglected. Mortality figures suggest successful

TABLE 43.3 Age-Standardized Mortality Rates by Cause for 2008

Region	Noncommunicable Disease	Communicable Disease	Cancer[a]	Cardiovascular Disease/Diabetes[a]	Chronic Respiratory Conditions[a]
Africa	779	798	147	382	92
Southeast Asia	676	334	125	322	109
Europe	532	51	166	238	17
Americas	455	72	136	169	24
Eastern Mediterranean	706	254	127	344	46
Western Pacific	534	74	168	184	41

[a] *Ages 30–70.*
Note: Data is per 100,000 population.
Source: WHO data [25,26].

treatment for these diseases in developed countries and little to no treatment in many parts of Africa. However, a degree of underreporting in many countries is likely.

Table 43.5 shows recent mortality data for HIV/AIDS, malaria, and tuberculosis (TB) for the major regions of the world. These diseases remain major problems in many African countries. In 2011, more than 20 million cases of malaria were reported in the African region. [26] Data for HIV/AIDS, malaria, cardiovascular disease, diabetes, and other causes of death indicate a wide diversity in prevalence and mortality between African countries. For example, 2011 data for HIV prevalence vary from <1% in Algeria to >25% in Swaziland. Death in 2010 due to malaria also shows great variation across African countries, with some reporting less than 1 death per 100,000 population and others reporting more than 150 per 100,000.

43.2.3 Population Diversity and Stratified Medicine

Population-based differences in disease susceptibilities and drug response were already observed before the era of genetics/genomics. The practice of stratified medicine therefore precedes pharmacogenetics/genomics. There has

TABLE 43.4 Cardiovascular Disease and Diabetes Mortality for Males by Region[a]

Region	Average Mortality Rate[b]	Lowest Rate (Country)	Highest Rate (Country)
Africa	448	279 (Algeria)	674 (Malawi)
US and Canada	171		
Latin America	289	148 (Peru)	545 (Trinidad & Tobago
Eastern Mediterranean	452	180 (Qatar)	765 (Afghanistan)
Western Europe (including Cyprus)	171	128 (France)	Cyprus (225)
Eastern Europe, Russian Federation	561	315 (Czech Republic)	881 (Turkmenistan)
South East Asia	392	343 (Thailand)	465 (Bhutan)
Western Pacific (Japan, Australia, Republic of Korea, New Zealand, Singapore)	153	118 (Japan)	171 (Singapore)
Western Pacific (excluding above)	475	293 (Brunei Darussalam)	922 (Nauru)

[a] *Average calculated from WHO 2008 global data for males.* [b] *Data per 100,000 population.*
Source: WHO data (age standardized estimated) [25,26].

TABLE 43.5 Deaths Due to HIV/AIDS, Malaria, and Tuberculosis

Region	HIV/AIDS, 2011[a]	Malaria, 2010[a]	TB Among HIV-Negative People, 2011[a]
Africa	139	72	26
Southeast Asia	12	2.4	26
Americas	9	0.2	2.2
Eastern Mediterranean	7.7	3.5	16
Europe	11	No data	5.0
Western Pacific	4.4	0.2	6.9

[a] *Data per 100,000 population.*
Source: Adapted from WHO [26,27].

been use of drugs based on patient race/ethnicity, but it is surrounded by serious controversy. With respect to people of African origin, the first example of this practice was in the United States when BiDil, a medication for congestive heart failure, was licensed specifically for use in American patients who self-identified as black [28]. Early studies had shown that African American patients with congestive heart failure generally responded less effectively to traditional treatments than did white patients with similar conditions [29]. Some populations have been associated with higher susceptibilities to some diseases; for example, Caucasians have a higher risk for coronary artery disease compared to African Americans and West Africans [30], while the opposite is true for hypertension [31].

43.3 POPULATION GENETICS, DIVERSITY, ANCESTRY, AND ADMIXTURE

Given the definition of population as encompassing race and ethnicity, specialist scientists have come up with further definitions; geneticists define it as a local group of people sharing a common gene pool and anthropologists define it as a group of people sharing a common language. In-depth discussion of these definitions is beyond the scope of this chapter; however, many studies observe an interesting overlap in the differently defined populations, which is not surprising given their shared geographical ancestry and genealogical relationships. The relevance of an understanding of Africa's population genetics and its diversity and admixtures lies in the potential use of such knowledge in addressing healthcare issues on the continent.

43.3.1 Genetic Variation and Diversity in African Populations

Genetic diversity, the level of biodiversity, refers to the total number of genetic characteristics in the genetic makeup of a species. It is distinguished from genetic variability, which describes the tendency of genetic characteristics to vary both within and among populations. Genetic diversity serves as a way for populations to adapt to changing environments. With more variation, it is more likely that some individuals will possess variant alleles that are suited for the environment. Those individuals are more likely to survive to produce offspring who bear that allele.

The determination of genetic diversity is based on several hypotheses and theories. These include the *neutral theory of evolution*, which proposes that diversity is the result of the accumulation of neutral substitutions; and the *diversifying selection theory*, which proposes that two subpopulations of a species that live in different environments select for different alleles at a particular locus. This may occur, for instance, if a species has a large range relative to the mobility of individuals within it. The *frequency-dependent selection theory* proposes that as alleles become more common, they become more vulnerable. This occurs in host–pathogen interactions, where a high frequency of a defensive allele in the host means that it is more likely that a pathogen will spread if it is able to overcome that allele [20].

The nature and importance of *human genetic variation*, on the other hand, address the differences between individuals and populations due to allelic variations. The allelic variations between individuals are a result of the exchange of genetic material during meiosis and various mutational events. Alleles occur at different frequencies in different human populations, with populations that are more geographically and ancestrally remote tending to differ more.

Genetic variation between populations is driven by the *natural selection* and *neutrality* of most mutations. Natural selection may confer an adaptive advantage to individuals in a specific environment if an allele provides a competitive advantage. Alleles under selection are likely to occur only in those geographic regions where they confer an advantage. The neutrality of most mutations encompasses the fact that they do not appear to have any selective effect on the organism one way or the other. In this case, the main cause of population differences is *genetic drift*, which is the effect of random changes in the gene pool. In humans, founder effect and past small population size (increasing the likelihood of genetic drift) may have an important influence on neutral differences between populations.

Based on these premises about population genetics, numerous studies have been conducted in populations from around the world, including those in Africa. Our overview of these starts by looking at the results of some of the latest major genetic studies done with African populations and makes reference to older studies that made similar findings. Emerging detailed data from genome-wide genotyping and whole-genome sequencing studies are highlighted.

The most comprehensive study of African populations so far, with respect to both the number of African and non-African populations studied and the genetic markers investigated, is that of Tishkoff et al. [5]. In this landmark work, the investigators conducted genetic analysis on 121 African populations, 4 African American populations, and 60 non-African populations for patterns of variation in 1327 nuclear microsatellite and insertion/deletion (ID) markers; this was the greatest number of markers used before the application of whole-genome microChip genotyping tools for the study of population structure and diversity.

43.3.2 High Genetic Diversity in African Populations

Genetic diversity was estimated by two methods [5], one based on the variance of microsatellite allele length (i.e., the microsatellite mutation rate) and the other based on heterozygosity (i.e., the proportion of individuals

TABLE 43.6 Inferred Population Clusters Using STRUCTURE Analysis of Autosomal Microsatellite and ID Polymorphism Data from Global Populations

K = cluster	Emerging Clusters
2	African,[a] non-African
3	East Asian, Oceanic, Native American
4	Eastern Africa[a]
5	Hadza,[a] Sandawe,[a] Southern African Khoisan (SAK),[a] Pygmy[a]
6	Western Pygmy[a]
7	Chadic,[a] Nilo-Saharan[a]
8	Indian, Oceanic
9	Oceanic
10	Native American
11	Mbuti Pygmy,[a] SAK[a]
12	Chadic/Nilo-Saharan[a] speakers from Cameroon, Chad,[a] Southern Sudan[a]
13	Sandawe[a]
14	Fulani[a]

[a] *African populations.*

Source: Data from Scheinfeldt et al. [32].

heterozygous at the loci studied). Using both measures, diversity was found to be high in African populations and higher than that found in European/Middle Eastern, Indian, East Asian, Oceanian, and Native American populations [5]. Genetic diversity also tended to decrease in proportion to the distance of a population from Africa, supporting what is known as the out-of-Africa model of human migration, in which serial founder effects are associated with reduced heterozygosity.

Further analysis of genetic diversity with respect to microsatellite variance and heterozygosity showed significant correlation with linguistic, geographical, and subsistence classifications. The hunter-gatherer populations (e.g., Pygmies, Bakola Pygmies, and San) were observed to have the highest genetic diversity. Phylogenetic trees constructed from genetic distances between populations also showed clustering by major geographic regions, both on a global scale and within Africa. These correlations support models of human population structure based on language classifications, archeological findings, geographical ocation, and livelihood. The genetic studies, however, applied both molecular and quantitative measures to these population structures.

43.3.2.1 Genetic Structure of African Populations

The population structure in Tishkoff's study [5] was explored using principal component analysis (PCA) and a Bayesian model-based approach with STRUCTURE software. Using STRUCTURE, the study inferred 14 population clusters (K1–K14). The African samples clustered primarily by geographical region and to a lesser extent by linguistic affiliation (Table 43.6).

The PCA of global populations revealed 72 significant components (PC1–72). These results tally with those obtained using STRUCTURE in that the first component, PC1 (accounting for 19.5% of the extracted variation) distinguished African and non-African populations. The admixture populations—African Americans and South African Cape Coloreds—clustered between the African and non-African population along PC1. Other PCs differentiated populations as indicated in Table 43.7. The PCA data can be easily visualized using the PCA plots in Figure 43.2(a,b).

The work by Tishkoff and colleagues [5] used 1327 nuclear microsatellite and ID markers to infer population structure and clustering. In 2008, we established a Biobank and Pharmacogenetics Database for African populations encompassing 10 ethnic groups from Kenya, Nigeria, Tanzania, and Zimbabwe [33]. Using the Illumina beadchips (www.illumina.com) and the genotype information of 364,648 SNPs, Maimbo and coworkers [34] characterized 467 unrelated individuals from these groups (Kenya: Maasai, Luo, Kikuyu, and a mixed population; Nigeria: Yoruba, Ibo, and Hausa; Tanzania: mixed populations; and Zimbabwe: Shona and San). Data on other African and non-African populations were taken from the HapMap project (www.hapmap.org/).

With PCA, we can discern genetic differentiation among all identified populations (Figure 43.3). The African populations were clearly distinguishable from Indian, Mexican, and Asian (Japanese and Chinese) populations and from those with European ancestry. In addition, African populations showed clustering patterns similar to those for the Yoruba (YRI), Luhya (LWK), and Maasai (MKK) in the HapMap project, compared with the other populations (Figure 43.3).

Population structure between the African ethnic groups was examined using FRAPPE analysis, which uses a maximum likelihood method to infer the genetic ancestry of each individual. A total of 98,196 SNPs were assessed by FRAPPE. The results were consistent with the those obtained by PCA (Figure 43.3(b)). At K = 2 up to K = 7, the data showed the same number of clusters as were observed in PCA, coinciding with the ancestral populations non-Bantu (Nigerian Yoruba, Ibo, and Nigerian Yoruba from HapMap); Bantu (Zimbabwean Shona and Tanzanian mixed

TABLE 43.7 Global Population PCA Analysis and within Africa's Population Structure

Principal Component	Percent Variance Explained	Populations Distinguished
Genetic structure on a global scale		
PC1	19.5	Distinguishes African and non-African. Mixed populations, African Americans, and Cape Mixed Ancestry clustered between African and non-African
PC2	5.01	Distinguishes Oceanians, East Asians, and Native Americans
PC3	3.5	Distinguishes Hadza hunter–gatherers
Genetic structure within Africa		
PC1	10.8	Distinguishes Saharan Africans from Western, Central and Southern Africans
PC2	6.1	Distinguishes Hadza
PC3	4.9	Distinguishes Pygmy and SAK

Source: Data from Tishkoff et al. [5].

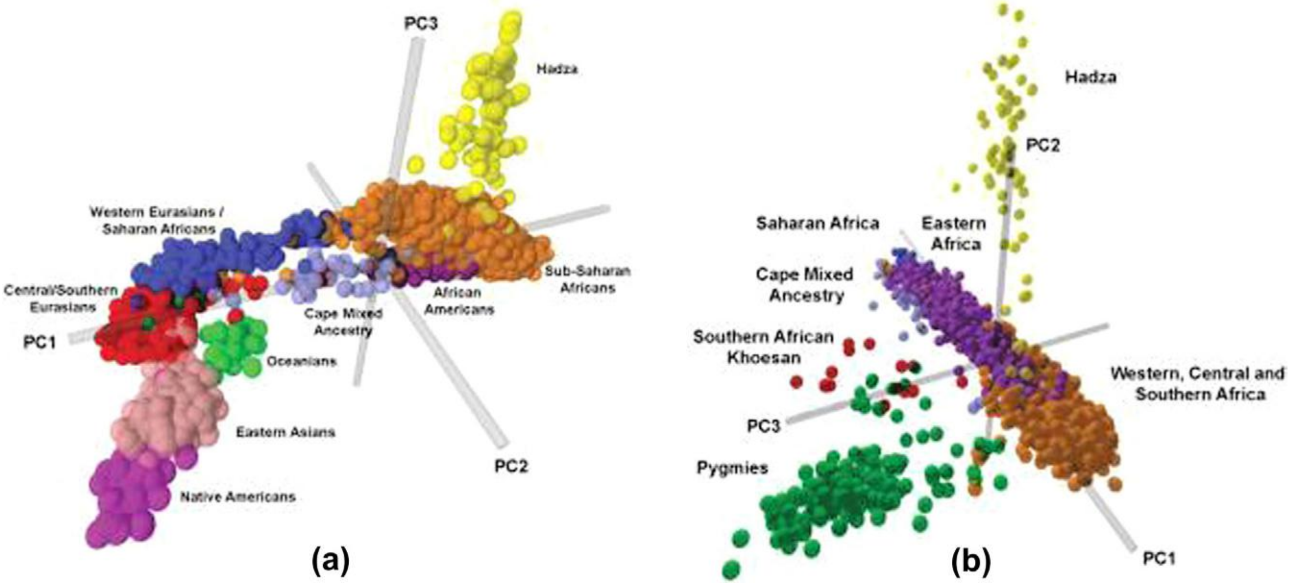


FIGURE 43.2 Principal components analysis for genetic structure at a global level (a) and within Africa (b). *Source: Reprinted with permission from Science: From Tishkoff et al.* [5].

populations), Afro-Asiatic (Nigerian Hausa), Khoisan (Zimbabwean San), and Nilotes (Kenyan Luo, Luhya, Maasai Kikuyu, and HapMap Maasai).

The Luhya, who have been reported as belonging to the Bantu (www.ethnologue.com/), were clearly clustered with the Nilotes, supporting the PCA data presented in this study (see Figure 43.3). Analysis at K = 5 revealed additional substructures in the Kenyan ethnic groups with Nilotic ancestral origin, dividing the group into three main clusters that are also similar to those observed in PCA. Bantu and non-Bantu populations were clearly differentiated up to K = 11, which result is similar to those reported by Bryc et al. [35].

F-statistics (F_{ST}) provides a measure of genetic structure. We have used average F_{ST} as a measure of genetic diversity to evaluate differences between populations calculated over all the autosomes (Table 43.8). The genetic diversities between and among African populations are generally larger than those observed between and among the Asian, Mexican,

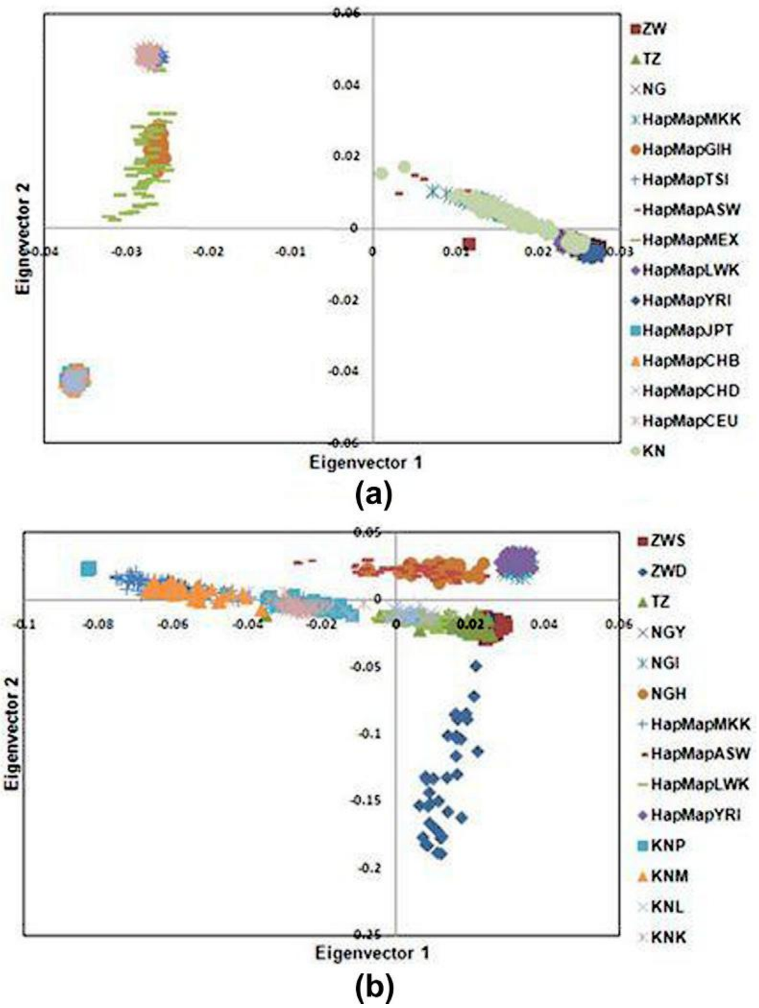


FIGURE 43.3 **PCA of world populations.** (a) Assemblage of African Biobank populations (Zimbabwe, Tanzania, Nigeria, Kenya) in relation to populations in the HapMap European (HapMap TIS, HapMap CEU), Mexican (HapMap MEX), Indian (HapMap GIH), Asians—Chinese (HapMap CHB, HapMap CHD) and Japanese (HapMap JPT), Kenya Luhya (HapMap LWK), Kenya Maasai (HapMap MKK), and Nigeria Yoruba (HapMap YRI) based on the two most informative principle components. (b) Using the two eigenvectors, assemblage of African HapMap populations: HapMap YRI, HapMap MKK, and HapMap LWK; African Biobank ethnic groups: Zimbabwe San (ZWD), Zimbabwe Shona (ZWS), Kenya Kikuyu (KNK), Luo (KNL), Maasai (KNM); Kenya mixed groups (KNP), Nigeria Ibo (NGI), Yoruba (NGY), Hausa (NGH); and mixed ethnic groups from Tanzania (TZ). *Source: Maimbo et al., unpublished data* [34].

Indian, and European populations. A low F_{ST} of 0.0393 was observed between the mixed Kenyans and individuals of West European ancestry, thus confirming the large genetic distance between these populations. Again average F_{ST} for individuals of African ancestry in the Southwest United States (ASW) from the HapMap data were as expected, showing that they are more similar to other African populations than to European and Asian populations. These data indicate that results from F_{ST} and PCA analysis are in agreement and agree with the geographical background of the populations under consideration (Figure 43.3(a) and Table 43.8).

43.3.2.2 *The Out-of-Africa Population Origin Model*

Most of the genetic data described so far point to human origin in Africa. Ancestral Africans maintained a large and subdivided population and experienced complex patterns of population expansion, contraction, migration, and admixture during their evolutionary history. The bottleneck associated with the founding of non-African populations (50–100 KYA) resulted in lower levels of genetic diversity, an increase in linkage disequilibrium (LD), and more similar LD patterns. Figure 43.4 schematically indicates the proposed out-of-Africa model, according to

TABLE 43.8 Genetic Differentiation between HapMap and African Biobank Populations

	CEU	CHB	CHD	GIH	JPT	LWK	MEX	MKK	TSI	YRI	KNK	KNL	KNM	KNP	NGH	NGI	NGY	TZ	ZWD	ZWS
ASW	0.0431	0.0652	0.0658	0.0431	0.0658	0.0086	0.0478	0.0081	0.0450	0.0078	0.0091	0.0103	0.0118	0.0099	0.0097	0.0097	0.0105	0.0105	0.0173	0.0090
CEU		0.0521	0.0526	0.0187	0.0531	0.0681	0.0161	0.0487	0.0042	0.0751	0.0474	0.0565	0.0437	0.0393	0.0561	0.0709	0.0643	0.0556	0.0614	0.0742
CHB			0.0032	0.0382	0.0062	0.0826	0.0348	0.0651	0.0536	0.0863	0.0683	0.0756	0.0659	0.0606	0.0754	0.0869	0.0825	0.0749	0.0810	0.0866
CHD				0.0386	0.0068	0.0830	0.0356	0.0654	0.0542	0.0865	0.0689	0.0762	0.0665	0.0613	0.0760	0.0873	0.0831	0.0756	0.0817	0.0868
GIH					0.0390	0.0631	0.0194	0.0441	0.0188	0.0684	0.0467	0.0554	0.0432	0.0399	0.0554	0.0674	0.0628	0.0552	0.0612	0.0681
JPT						0.0833	0.0352	0.0661	0.0546	0.0871	0.0688	0.0760	0.0664	0.0610	0.0757	0.0874	0.0829	0.0753	0.0814	0.0873
LWK							0.0632	0.0098	0.0677	0.0061	0.0069	0.0044	0.0111	0.0075	0.0064	0.0069	0.0071	0.0054	0.0116	0.0049
MEX								0.0402	0.0180	0.0649	0.0537	0.0645	0.0489	0.0499	0.0650	0.0714	0.0715	0.0659	0.0720	0.0661
MKK									0.0455	0.0147	0.0048	0.0077	0.0034	0.0045	0.0090	0.0137	0.0123	0.0093	0.0137	0.0138
TSI										0.0735	0.0490	0.0594	0.0445	0.0418	0.0592	0.0723	0.0674	0.0591	0.0651	0.0731
YRI											0.0093	0.0062	0.0145	0.0092	0.0047	0.0031	0.0032	0.0054	0.0112	0.0049
KNK												0.0086	0.0078	0.0075	0.0105	0.0116	0.0127	0.0101	0.0168	0.0086
KNL													0.0121	0.0103	0.0088	0.0083	0.0096	0.0081	0.0152	0.0055
KNM														0.0084	0.0141	0.0172	0.0179	0.0150	0.0210	0.0143
KNP															0.0123	0.0124	0.0144	0.0124	0.0189	0.0089
NGH																0.0068	0.0079	0.0094	0.0166	0.0061
NGI																	0.0049	0.0074	0.0141	0.0053
NGY																		0.0089	0.0160	0.0056
TZ																			0.0146	0.0042
ZWD																				0.0099

[a] Average F_{ST} between two populations is shown. The HapMap populations are European (TIS, CEU), Mexican (MEX), Indian (GIH), Asians (Chinese CHB, CHD), Japanese (JPT), Kenya Luhya (LWK), Kenya Maasai (MKK), and Nigeria Yoruba (YRI). The African Biobank populations are Zimbabwe San (ZWD), Zimbabwe Shona (ZWS), Kenya Kikuyu (KNK), Luo (KNL), Maasai (KNM), Kenya mixed groups (KNP), Nigeria Ibo (NGI), Yoruba (NGY), Hausa (NGH), and Tanzanian mixed ethnic groups from (TZ).

Source: Maimbo et al. [34]

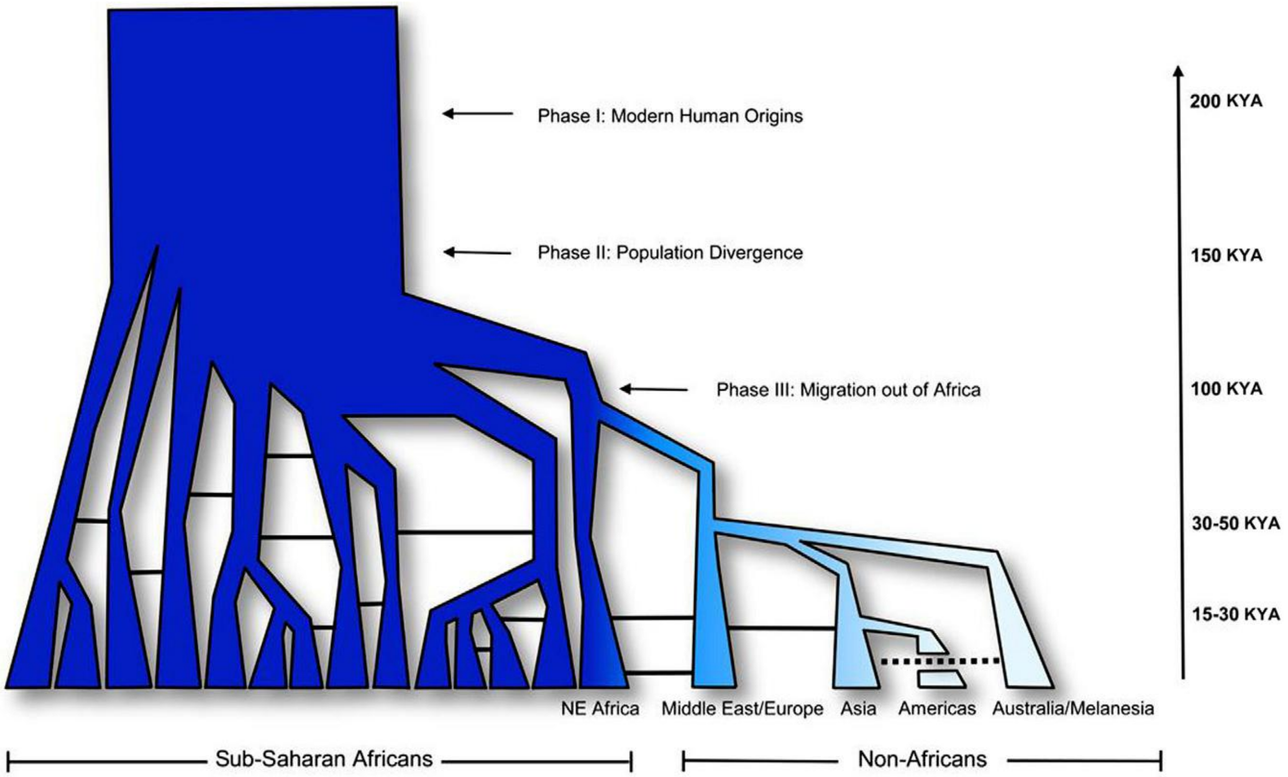


FIGURE 43.4 **Origin of modern humans, population divergence, migration out of Africa, and changes in LD due to founder effects.** Solid horizontal lines indicate gene flow between populations. *Source: Reprinted with permission from Elsevier: Data from Campbell and Tishkoff* [2].

which, anatomically modern humans evolved in Africa around 200 KYA, migrated to Euroasia within the last 40–80 KYA, and then migrated to the Americas within the last 15–30 KYA.

43.3.2.3 African Populations' Admixture Within and Outside of Africa

The history of Africa has been characterized by both ancient and recent events that have resulted in population admixtures. Some of these are visually easy to discern, as in populations resulting from the mixture of African and non-African populations; some are not visible, however, and require genetic analysis to uncover. African Americans and South African Cape Mixed Coloreds represent established admixture populations, whereas recent global population movements are giving birth to more complex admixture patterns.

The slave trade of the eighteenth century (1711–1800) is estimated to have resulted in the shipment of more than 40,000 people from the West African coast (Senegal to Cameroon) [36]. In the study carried out by Tishkoff and colleagues [5], African Americans from Chicago, Baltimore, Pittsburgh, and North Carolina were found to have substantial ancestry from the African Niger-Kordofanian and from Europeans. European and African ancestry varies considerably in African Americans, as later shown in detail by Bryc and coworkers, [35] who analyzed African Americans for genomic admixture by PCA using the Affymetrics GeneChip 500 K, and showed European ancestry to be on average 18.5% but with very large interindividual variation. The X-chromosome showed elevated levels of African ancestry, consistent with a sex-biased pattern of gene flow, with an excess of European male and African female ancestry. Levels of African and European genetic admixture have also been observed in genes coding for drug-metabolizing enzymes.

The Cape Mixed Ancestry population of South Africa showed the highest level of intercontinental admixture of any global population, with nearly equal high levels of Southern African Khoisan (SAK), Niger-Kordofanian, Indian, and European ancestry. This level of admixture has been shown to have important implications in the variability of genes responsible for drug disposition.

43.3.2.4 Dietary Factors in Population Diversity and Evolution

Natural selection and random genetic drift are two factors that may affect allele frequency in a population. Both may lead to the elimination or the preferential selection of a particular allele. Evolution of the *LCT* gene in world populations, and especially in Africans, has been well documented [37]. This gene is the determinant of the trait commonly called lactase persistence/nonpersistence, depending on whether lactase activity continues from childhood into adulthood.

Populations in various parts of Africa began domesticating animals as a source of food and for ploughing more than 7500–9000 years ago in southern Egypt [38,39]. The foundation of the selection based on *LCT* variation is that individuals raised in areas where dairying is commonly practiced maintain lactase activity into adulthood while individuals in areas where there is less dairying become lactase nonpersistent as adults. Tishkoff and colleagues [37] looked at the convergent adaptation of human lactase persistence in Africa and Europe and found that the frequency of lactase persistence was lowest in the Khoisan-speaking Sandawe hunter-gatherer population from Tanzania (26%).

In our genome-wide genotyping study [34], among the most highly differentiated SNPs, rs12472293, rs1050115, and rs961360 were found in the surrounding area (within 250 kb) of an LD block containing *LCT* and the *minichromosome maintenance gene 6* (*MCM6*). We observed that the Kenya Kikuyu (KNK) and the Nigeria Hausa (NGH) maintain minor allele frequencies (MAFs) for all three SNPs in the *LCT* gene at ~30%. On the other hand, the MAFs in the San population were observed to be less than 20%.

Genetic variation has also been shown to occur in genes coding for proteins important in disease susceptibility. Such variation has been associated with selection of certain genotypes that confer resistance to some infections. In Africa, important susceptibility and/or resistance genes have been documented for HIV/AIDS, tuberculosis, and malaria.

The next section focuses on pharmacogenomics, which is the study of genetic variation in genes that influence drug disposition and efficacy, and its implications for stratified medicine.

43.4 PHARMACOGENOMICS IN AFRICA

Pharmacogenomics refers to the study of total genetic variation in an individual that affects drug response. Although sometimes pharmacogenomics is used interchangeably with pharmacogenetics, the latter refers to studies on a specific gene or a few genes that are known to be rate-limiting in the metabolism of a particular chemical or drug. Many studies to date have been of a pharmacogenetic nature, focusing on genes involved in the pharmacokinetics (absorption, distribution, metabolism, and excretion) and the pharmacodynamics (mechanisms of action) of drugs [40,41]. Post-human genome sequencing has seen an exponential increase in high-throughput genomic technologies that have enabled mass identification of pharmacogenes at a faster pace and less expensively. These technologies have included RNA expression arrays, genome-wide association studies (GWAS) and targeted whole-exome or whole-genome sequencing.

Such advances have enabled evaluation of genes showing significant changes in expression as well as identification of single-nucleotide polymorphism (SNP) profiles for patients with different drug response patterns. The myriad ongoing studies of variation in the human genome have seen the curation of gene variations associated with drug response into The Pharmacogenomics Knowledge Base (PharmGKB, www.pharmgkb.org). This database was developed by the US PharmacoGenetics Research Network (PGRN) to aid researchers in investigating how genetic variation affects drug response [42]. However, pharmacogenomic determinants of drug response differ between as well as within populations. Africa remains understudied with respect to genomic variation. This section evaluates the pharmacogenomics studies that have been carried out among African populations.

Before the sequencing of the human genome was completed (pre-2000), pharmacogenomic studies in African populations were largely focused on candidate SNPs or candidate gene analysis, replicating earlier reports on either Caucasian or Asian populations. There was lack of appreciation of the inherent qualitative and quantitative differences in genetic variants between different populations to the extent that African populations were treated as homogenous, with data observed in one ethnic group often extrapolated to represent all African groups.

Most current drug discovery and development are carried out on a genetic backbone of either Caucasian or Asian populations, when a new drug is prescribed for African populations, it may possibly be the first time it is exposed to some of Africa's unique genetic variations. This may result in unexpected ADRs. For example, abacavir, a nucleoside reverse transcriptase inhibitor (NRTI) used in the treatment of HIV-1, results in hypersensitivity in nearly 8% of patients carrying the *HLA-B*5701* allele. This variant is rare in African populations, however, and relatively high in Indian populations (up to 20%), which shows that genetic testing for this HLA variant across all world populations may not be useful [43].

In these early stages of the post–genome sequencing era, publications on pharmacogenomics with an African connection have mostly reported on candidate genes or targeted gene domain sequencing, including both nuclear and mitochondrial genomes [44–53].

43.4.1 Pharmacogenetic Diversity in African Populations

The genetic diversity of African populations has been well demonstrated, as described in Section 43.3, through the use of various genetic markers, including genome-wide genotyping, microarrays for more than 500 K SNPs, microsatellite markers, and mitochondrial DNA. We turn now to the pharmacogenetic diversity that exists between Africa and the rest of the world, as well as the diversity that exists between African populations themselves.

Studies of genes encoding drug-metabolizing enzymes (DMEs), such as the cytochrome P450 (CYP450) and glutathione S-transferase superfamilies, have identified variations that affect the use and efficacy of therapeutic drugs. CYP enzymes are responsible for the phase 1 hepatic and intestinal metabolism of a wide spectrum of endogenous and xenobiotic compounds. The most clinically relevant CYP enzymes are CYP1A2, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP3A4, and CYP3A5, and they are thought to metabolize more than 90% of all therapeutic drugs [54–56]. CYPs metabolize a broad spectrum of substrates; thus, functional variation in their genes is associated with individual and population differences that have pharmacogenomics relevance, especially with respect to adverse drug events (ADEs).

Early studies of drug-handling variation in African and other populations were of a phenotyping nature, especially for CYP2D6, CYP2C19, and NAT2 activities. CYP2D6 was the most studied enzyme, using phenotyping probe drugs (e.g., debrisoquine, dextromethorphan, metoprolol, codeine, and sparteine), while CYP2C19 and NAT2 were phenotyped using mephenytoin and caffeine, respectively. Metabolic ratios—parent drug conc/metabolite(s)—were used, and cut-off points were designated (antimodes obtained from probit analysis of population metabolic ratio data) for classifying individuals broadly into poor and extensive metabolizers (PM and EM) or fast and slow acetylators.

Later studies characterized the genetic basis for the phenotypes and launched numerous studies in different populations linking phenotype with genotype. PMs are usually homozygous or heterozygous for null alleles, whereas EMs comprise a wide spectrum of individuals including normal allele homozygotes, heterozygotes with at least one functional allele regardless of level of activity, and homo- or heterozygotes of reduced functional activity. Recently, therapeutic (dose) recommendations were published [57] for gene-drug combinations comprising 53 drugs and 11 genes. For example, for doxepin, it is recommended that CYP2D6 PMs receive 60% reduced dosage, while ultrarapid metabolisers (UM) should receive a 100% dose escalation.

Table 43.9 and Figure 43.5 provide a comprehensive analysis of a selected number of SNPs in genes coding for drug-metabolizing enzymes in many African and non-African populations. PCA analysis results clearly show the clustering of the major world populations—Asians, Caucasians, and Africans—in the first two PCs (PC1 and PC2; Figure 43.5). This is consistent with the population clustering found by Maimbo et al. [34] using markers, shown in Figures 43.2 and 43.3.

An excellent review on the phenotype and genotype of CYP polymorphisms in African populations has been written by Alessandrini and coworkers [58]. In it the authors demonstrate the genetic diversity of CYP450 alleles among the various African populations that can have clinical implications. This further supports previous work by Matimba and colleagues [33] who conducted extensive population-genotyping studies with African populations.

The variability in genes important for drug response has significant implications for the safe and efficacious use of drugs, and it necessitates the inclusion of this genetic information in the discovery and development of new chemical entities for treatment of disease. Such knowledge can also inform and guide the design of clinical studies that may produce population- and regional-specific product labeling.

Although the burden of infectious and parasitic diseases is great in African populations (e.g., HIV/AIDS, TB, and malaria), early studies in pharmacogenetics among Africans focused mostly on genes whose enzymes are involved in the metabolism of drugs used in noncommunicable diseases. For example, studies have been carried out on CYP2D6 and NAT2, the relevance of which is mostly for antidepressants, antipsychotics, and anticancer drugs, respectively. Despite the weakness in early pharmacogenetics studies in African populations, a substantial amount of information has begun to accumulate on the genetic diversity of Africans, especially with respect to genes of pharmacogenetic importance. Further fine genetic differences between ethnic groups are being revealed as more populations are studied [59,60].

There is much more pharmacogenetic data on CYP2D6 with respect to Africans and other populations than any other pharmacogene because of its involvement in the metabolism of commonly used drugs, most of which have narrow therapeutic indices (e.g., antidepressants and antipsychotics). In populations of European descent, phenotyping and genotyping established the main CYP2D6 genetic variants that can explain the CYP2D6 phenotype; these include *CYP2D6*2*, **3*, **4*, and **5*. These variants constitute genotypes that result in good correlations with phenotypes and are naturally replicated in other populations, including African populations.

Pioneering work in African populations on the pharmacogenetics of CYP2D6 and CYP2C19 was carried out among Zimbabweans [61–65], Ghanaians [44,66], Tanzanians [67,68], Ethiopians [69], Zambians [70,71], and South African Mixed Coloreds and Xhosa populations [72–74]. New variants were discovered and characterized (e.g., *CYP2D6*17* [63]) and their functional effects on drug metabolism were subsequently characterized [75,76]. In the study of CYP2D6*17 function in Tanzanians, who were genotyped and phenotyped for CYP2D6, it was observed that all carriers of the *CYP2D6*5/*17* genotype are EMs with codeine yet are all PMs with respect to debrisoquine and dextromethorphan [75]. From this and other studies, it became obvious that classification of people into EM, UM, and PM with respect to CYP2D6 in African populations is probe drug–dependent. In addition, frequencies of each variant in different populations were shown to differ; for example, among Ethiopians, *CYP2D6 *2N* is the most common allele instead of *CYP2D6*6*, as in Caucasians [69].

TABLE 43.9 Allele Frequencies (%) of DME Genetic Polymorphisms

Populations	CYP2C19		CYP2D6								NAT2				GST		CYP2B6	FMO		
	*2	*3	*2/2	*3	*4	*5	*10	*9	*17	*29	*5	*6	*7	*14	M1 del/del	T1 del/del	*6	K158	M257	G308
Orientals	30	10	2	0	1	6	51	0	0	0	5	25	13	0	55	65	18	-	-	-
Chinese	37	8	1	0	1	6	51	0	0	0	6	31	16	0	58	53	21	22.9	20.3	14.8
Japanese	35	11	1	0	1	3	43	0	0	0	2	19	10	0	44	44	16	22.7	14.5	21
Koreans	21	12	0	0	2	6	51	0	0	0	3	19	11	0	53	60	15	18.9	-	18.3
Caucasians	15	0	5	2	25	5	2	2	0	0	49	27	2	0	50	15	21	-	-	-
Swedes	17	0	1	3	23	5	1	0	0	0	51	28	2	0	51	20	-	44.3	7.1	22.4
Germans	18	0	2	2	20	2	2	0	0	0	46	27	4	0	51	21	-	42.6	6.9	22.5
Americans	14	0	2	-	-	-	-	-	0	-	45	28	2	0	54	15	-	39-53	6.7-22	15-22
Mixed Africans	16	1	2	0	2	4	6	0	30	15	34	20	5	13	30	-	40	-	-	-
African American	25	0	1	0	7	6	4	1	15	5	30	22	2	9	28	24	47	41-52	7-May	0-5.2
Tanzanians	18	1	3	0	2	4	4	0	18	20	34	21	3	13	33	25	39	-	-	-
Shona	13	0	2	0	2	4	6	0	34	17	31	21	6	14	24	26	38	50	2	1.5
Venda	21	0	-	0	3	5	12	0	24	6	39	22	5	11	23	20	36	48.4	2.4	1.6
Ghanaians	-	-	2	0	7	6	3	0	28	-	-	-	-	-	39	-	49	-	-	-
Ethiopians	14	2	15	0	4	3	9	0	9	-	-	-	-	-	-	-	-	-	-	-
Kikuyu	16	0	-	0	1	-	-	0	33	14	58	24	-	-	28	25	34	49	4.6	1
Luo	18	0	-	0	4	-	6	0	23	16	34	22	3	14	29	22	37	50	4.5	0
Maasai	11	0	-	0	8	-	5	0	18	8	42	27	4	9	16	40	35	42	4.2	0
Igbo	29	0	-	0	8	-	10	0	14	20	28	29	4	11	23	36	38	43.9	0.5	0.5
Yoruba	10	0	-	0	3	-	7	0	22	10	33	27	3	8	31	35	42	52	1	0.5
Hausa	12	0	-	0	2	-	13	0	18	10	27	33	3	3	37	42	42	44	4	0.5
San	12	-	-	-	9	-	-	0	22	2	20	8	-	-	45	-	40	33.3	0	0.8

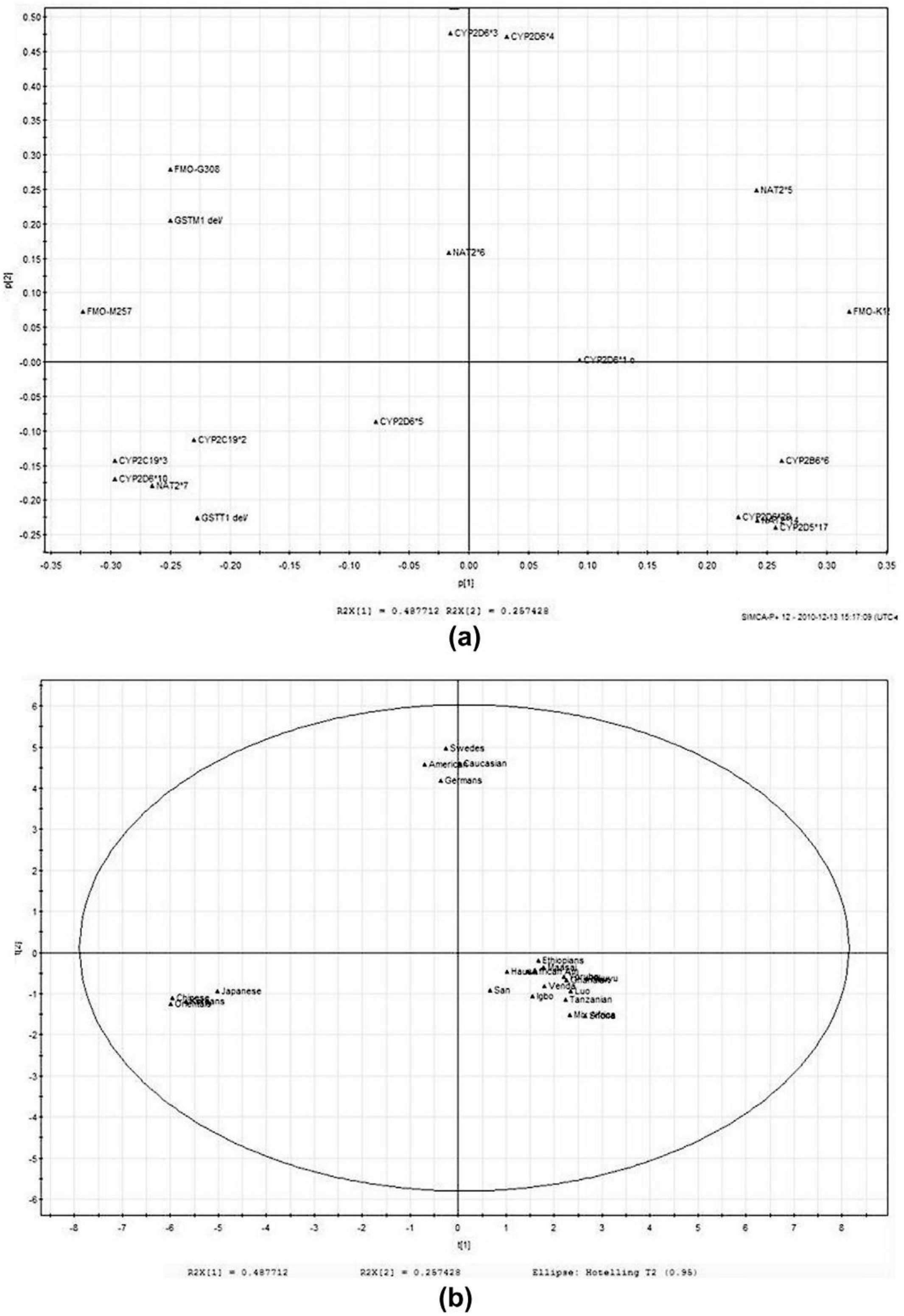


FIGURE 43.5 **PCA analysis of allele frequencies in Asian, African, and Caucasian populations.** (a) Loading plot showing correlations between SNPs; (b) score plot showing correlations between populations. *Source: Reproduced with permission from Nature Publishing Group, Eur J Hum Genetics. From Matimba et al.* [33].

43.4.1.1 *Variation in Pharmacogenetically Relevant Genes in African Populations*

In the following subsections, we discuss some very important pharmacogenes (VIPs)—*CYP3A5*, *CYP2C9*, *VKORC1*, *CYP2B6*, *NAT-2*, and *CYP2D6*—and their function in the metabolism of drugs relevant in Africa.

CYP3A5

CYP3A5 variability in protein expression is largely attributed to four alleles: *CYP3A5*1*, *CYP3A5*3*, *CYP3A5*6*, and *CYP3A5*7*, of which only *CYP3A5*1* is associated with CYP3A5 expression; the other three are nonexpressors. The *CYP3A5*3* allele varies in frequency between different populations, from as low as 4% in Africans to as high as 94% in Caucasians [56]. The frequency of the *CYP3A5*3/*3* homozygous genotype, which is associated with null CYP3A5 activity, differs among African populations, from as low as 0 to as high as 53%, while there are differences between ethnic groups in the same country—for example, among Cameroonian ethnic groups, the homozygous *CYP3A5*3/*3* genotype varies from 0 to 17% [56].

There are also qualitative differences when African populations are compared to other world populations. The *CYP3A5*6* and *CYP3A5*7* alleles are observed almost exclusively in individuals with recent African ancestry. The significance of the distribution of the various *CYP3A5* allelic variants seems to point to environmental influence. For example, Thompson and colleagues [77] reported that elevated *CYP3A5*3* frequencies tended to be positively correlated with increased geographic distance from the equator.

Bains and coworkers [56] reviewed evidence suggesting a strong positive correlation between latitude and functionally important variants of genes implicated in salt-sensitive hypertension. *CYP3A5* is involved in the metabolism of renal cortisol to 6-β-hydroxycortisol, which is a key regulator of renal sodium transport, and immune responses that cause inflammation; thus, *CYP3A5*1* provides a selective advantage in equatorial populations because of the role of CYP3A5 in salt retention and the reabsorption of water. On the other hand, it is detrimental to carry the *CYP3A5*3* allele in the same region. This is supported by the generally low frequencies of homozygous *CYP3A5*3/*3* in African populations around the equator—Ghana (up to 0.04), Senegal (up to 0.09), Malawi (up to 0.05), Nigeria (up to 0.02)—when compared to populations far from the equator—Ethiopia (up to 0.53) and most European populations (from 0.84 up to 0.90).

The frequency of *CYP3A5*3* in world populations is shown in Figure 43.6, which was taken from Bains and coworkers [56]. The geographical distribution of CYP3A5 variants shown has pharmacogenomic relevance for African populations because CYP3A5 is involved in the metabolism of drugs used to treat the three most common infectious diseases: malaria (artemether, lumefantrine, mefloquine, primaquine, chloroquine), HIV/AIDS (efavirenz, saquinavir, maraviroc, indinavir, nelfinavir, ritonavir, lopinavir), and tuberculosis (rifampicin). In a study among Ugandans [78], coadministration of artemether/lumefantrine with efavirenz was associated with reduced artemether levels.

CYP2B6 and Efavirenz Pharmacogenetics

The availability of ARV drugs has transformed HIV/AIDS from a "death sentence" to a chronic disorder. However, this chronic nature implies that HIV/AIDS patients are required to take medications for the rest of their lives. Most, if not all, of the available ARV drugs were developed in clinical trials that did not include African populations, and some have been shown to be unsuitable and lead to ADRs in a proportion of African patients [79]. Certain of these ADRs are serious and have led to patient noncompliance. ARV noncompliance is associated with development of virus resistance and the need to move patients to very expensive second-line ARV regimens. Most ARV drugs are substrates of at least one of the following CYPs: 1A1, 1A2, 1B1, 2A6, 2B6 2C8, 2C9, 2C19, 2D6, 2E1, 3A4, and 3A5; thus, they are subject to interpatient variability in plasma levels due to polymorphisms in genes coding for these enzymes. The combined antiretroviral drug therapy is commonly referred to as highly active antiretroviral therapy (HAART).

Efavirenz (EFV) forms the backbone of first- and second-line HAART regimens. EFV is primarily metabolized by CYP2B6 and to a minor extent by CYP1A2, CYP2A6, and CYP3A5 [80]. EFV metabolism by CYP2B6 leads to 8-hydroxy-EFV (8-OH-EFV), which accounts for 92% of the metabolic pathway, while some minor metabolites are produced through other pathways. There is high interindividual variability in plasma EFV levels after drug administration at the usual dosage regimen of 600 mg/day. Above-normal levels of EFV are associated with adverse effects such as neuropsychiatric manifestations. The SNP 516G>T is associated with diminished activity of the CYP2B6 isoenzyme, increased plasma EFV concentrations, and increased incidence of EFV-associated neuropsychological toxicity.

The preferable mid-dosing plasma level of EFV is 1–4 μg/ml to allow for optimized antiretroviral potency and to minimize the risk of toxicity. HIV-1 patients who receive standard doses and have a plasma EFV concentration of <1 μg/ml appear to have a higher risk for virological failure and emergence of selective drug resistance, while those with high plasma EFV concentrations of >4 μg/ml may experience adverse CNS effects more frequently.

Gounden et al. [81] found a correlation between *CYP2B6* 516T/T genotype, EFV concentrations, and early neuropsychiatric side effects. Several studies have confirmed these findings in African populations from Zimbabwe [82], Uganda [53], South Africa [49], and among

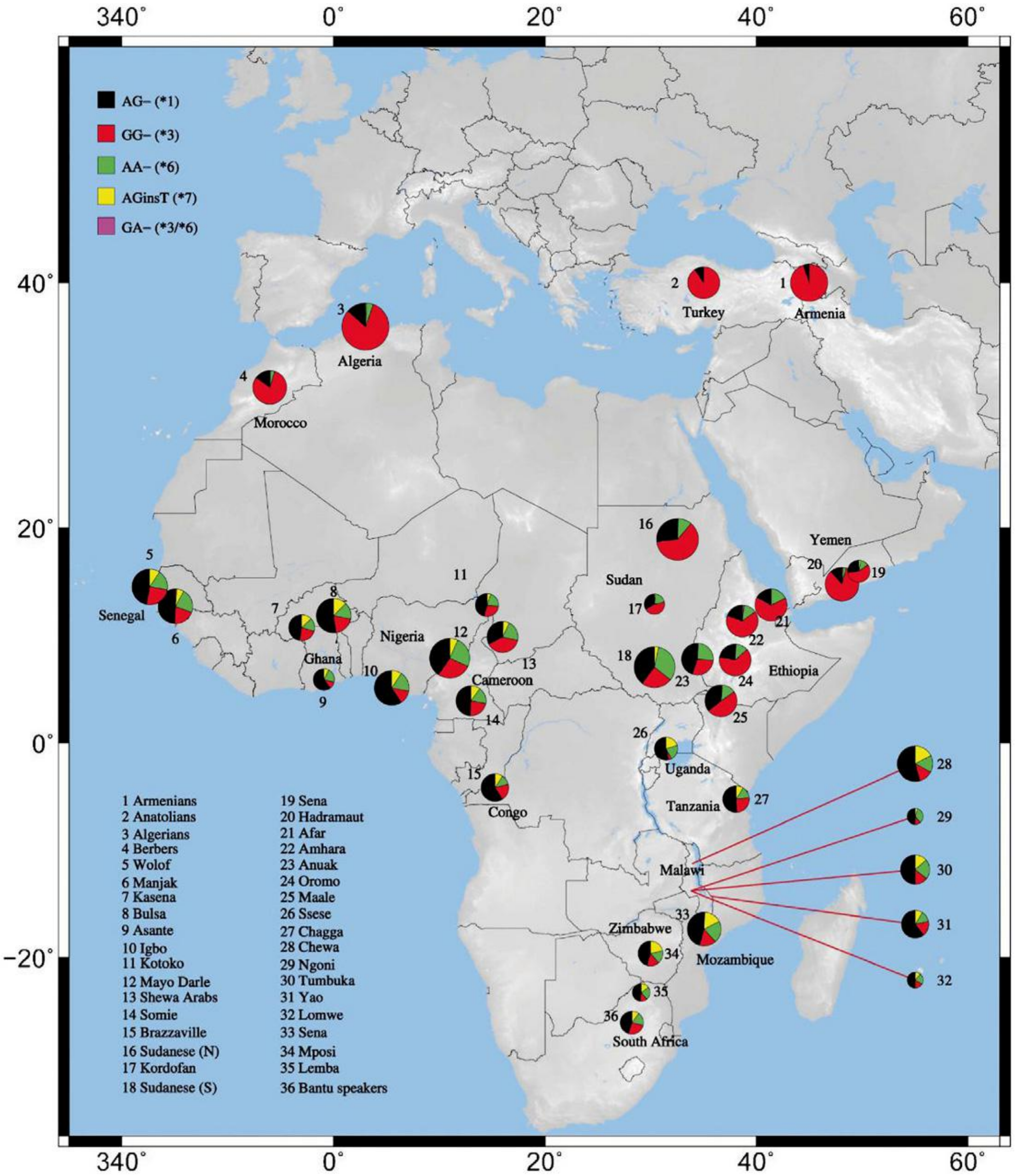

340°
0°
20°
40°
60°
40°
20°
0°
-20°
AG– (*1)
GG– (*3)
AA– (*6)
AGinsT (*7)
GA– (*3/*6)
1 Armenia
2 Turkey
3 Algeria
4 Morocco
5 Senegal 6
7 8 Ghana 9
10 Nigeria 11 12 13 Cameroon 14
15 Congo
16 Sudan 17 18
Yemen 20 19
21 22 23 24 Ethiopia 25
26 Uganda
Tanzania 27
28 29 30 31 32
Malawi
33 Zimbabwe Mozambique 34
35 36 South Africa
1 Armenians
2 Anatolians
3 Algerians
4 Berbers
5 Wolof
6 Manjak
7 Kasena
8 Bulsa
9 Asante
10 Igbo
11 Kotoko
12 Mayo Darle
13 Shewa Arabs
14 Somie
15 Brazzaville
16 Sudanese (N)
17 Kordofan
18 Sudanese (S)
19 Sena
20 Hadramaut
21 Afar
22 Amhara
23 Anuak
24 Oromo
25 Maale
26 Ssese
27 Chagga
28 Chewa
29 Ngoni
30 Tumbuka
31 Yao
32 Lomwe
33 Sena
34 Mposi
35 Lemba
36 Bantu speakers

FIGURE 43.6 **Distribution of the five inferred CYP3A5 haplotypes for some African populations.** Genotyping was performed in 2245 individuals from 32 geographically and ethnically diverse African populations for common clinically relevant CYP3A5 alleles; also genotyped were an additional non-African 293 individuals from Europe and the Arabian Peninsula. The size of each circle is proportional to the number of individuals sampled from a given population. The allele combinations at all three loci are given in the key. The alleles in brackets define the inferred haplotype. The lowest *CYP3A5*3* frequencies are found in Sub-Saharan Africa, although frequencies are higher in East Africa than in other populations sampled from the continent. *Source: Reprinted with permission from BMC Genetics: From Bains* [56].

Tanzanians [47]. Moreover, other studies have established that *CYP2B6* T983C increases the predictive capability of the *CYP2B6* G516T SNP for EFV pharmacokinetics [49,83]. Associations between increased plasma EFV exposure and *CYP2B6* 516G>T and 983T>C have been consistent across multiple studies and populations, moving us closer to the possible design of a pharmacogenomics diagnostic tool to aid in EFV-dosing decisions [49]. *CYP2B6*6* is more frequent in African populations compared to Caucasians and Asian populations (Table 43.9).

Studies to determine EFV dosing in Caucasians arrived at a dose of 600 mg/day. Population pharmacokinetic studies in African populations (Tanzania, Ethiopia, Zimbabwe, South Africa) showed that more than 50% of patients have EFV levels above the minimum safe level of 4 μg/ml. Pharmacometric modeling indicates a need to reduce the

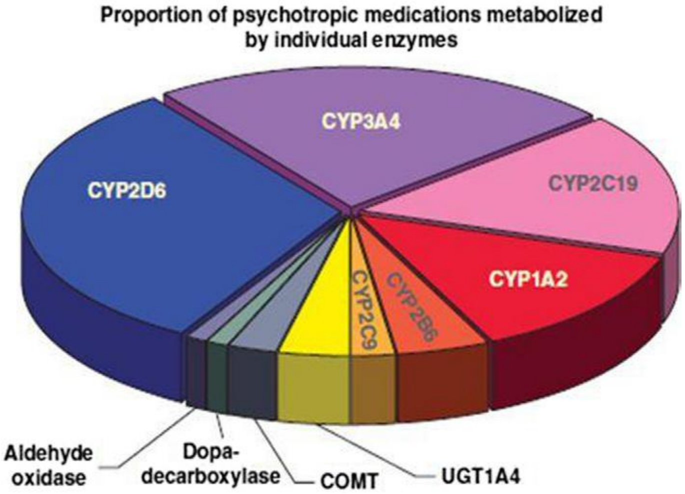


FIGURE 43.7 **Proportion of psychotropic medications metabolized by individual enzymes.** *Source: Reprinted with permission from Nature Publishing Group, Mol Psychiatry: From Stingl et al.* [84].

dose of EFV in a significant number of African patients to as low as 200 mg [79]. The modeling shows that patients homozygous for the *CYP2B6*6* require 200 mg, while women heterozygous for *CYP2B6*6* and weighing below 62 kg require 400 mg [79]. Studies in patients in South Africa also demonstrated that it is possible to predict patients likely to have high plasma concentrations based on the *CYP2B6*6* genotype [49]. While *CYP2B6*6* and some biometric measures explain about 22% of elevated EFV plasma concentrations [79], *CYP2B6*6*'s use as a biomarker for CNS ADRs still needs further investigation.

CYP2D6 and the Metabolism of Antidepressants and Antipsychotics

One of the most studied pharmacogenes, *CYP2D6*, has resulted in the development of genotyping kits such as the amplichip CYP450 by Roche, which is used to categorize CYP2D6 phenotypic activity based on CYP2D6 genotypes. These genotypes are computed from nearly 100 allelic variants that have been reported to date. With respect to CYP2D6, individuals are categorized as ultrarapid, extensive, intermediate, and poor metabolizers. As mentioned earlier, CYP2D6 phenotypes and genotypes show both qualitative and quantitative differences in world populations, including different activity levels for drugs that need CYP2D6 activation. Some of the substrates of CYP2D6 include the following: tamoxifen, sertraline, imipramine (Tofranil), doxepin (Adapin, Silenor, Sinequan), maprotiline (Deprilept, Ludiomil, Psymion), trimipramine (Surmontil), desipramine (Noraprim), nortriptyline (Aventyl, Pamelor), clomipramine (Anafranil), paroxetine (Paxil), venlafaxine (Effexor), amitriptyline (Elavil), mianserin, Trazodone (Desyrel), bupropion (Wellbutrin), nefazodone (Serzone), citalopram (Celexa) and sertraline (Zoloft); and antipsychotics: perphenazine (Trilafon), thioridazine (Mellaril), olanzapine (Zyprexa), zuclopenthixol (Cisordinol, Clopixol, Acuphase), aripiprazole (Abilify), flupentixol (Depixol, Fluanxol), haloperidol (Haldol), perazine (Taxilan), risperidone (Risperdal), and pimozide (Orap) [41].

Most drugs are metabolized by more than one enzyme, however, and Figure 43.7 illustrates this for psychotropic drugs generally. In addition to genetic polymorphisms, several inhibitors also compromise the effectiveness of substrate drugs by reducing CYP2D6's activity to varying extents. As an example of a CYP2D6 substrate, tamoxifen, prescribed for the treatment of estrogen receptor–positive breast cancer, is converted by CYP2D6 to endoxifen, the active metabolite. With CYP2D6 PM status in world populations ranging from 0.5% to nearly 10%, there are differences in the effectiveness of tamoxifen interindividually and between different populations. Besides CYP2D6 variations, there are genes whose expressed enzymes participate in tamoxifen metabolic and elimination pathways (e.g., CYP3A4, CYP3A5, SULT1A1, and UGT2B15); these need to be taken into account when designing a tamoxifen pharmacogenomic testing tool.

Hicks and colleagues [85] published "Clinical Pharmacogenetics Implementation Consortium Guideline for *CYP2D6* and *CYP2C19* Genotypes and Dosing of Tricyclic Antidepressants" to guide clinicians in the use of these drugs in Caucasians. These guidelines clearly do not address the case for African populations for whom African-specific variants for diminished CYP2D6 enzyme activity have been discovered—for example, *CYP2D6*17* [64] and *CYP2D6*29* [86]. No clinical studies have been performed to evaluate these variants' likely effects in patients on CYP2D6 substrate drugs, although Masimirembwa and Hasler [59] made postulations on their possible effect.

CYP2C9 and VKCOR1 the Disposition of Warfarin

Warfarin, a drug used for the long-term treatment and prevention of thromboembolic events, exhibits a narrow therapeutic window and shows substantial interpatient dose variability. It is one of the most used anticoagulants in the world. It is very difficult to determine the appropriate warfarin dose; thus, in the first few days of treatment, this drug is associated with one of the highest rates of ADEs. It is estimated that 20% of patients started on warfarin are hospitalized for ADEs during the first 6 months of treatment. The S-warfarin isomer, which is the more potent, is metabolized principally by *CYP2C9*; thus, its polymorphic variation affects the relative levels of circulating warfarin and other drugs [87].

Many allelic variants have been reported for *CYP2C9*, with *CYP2C9*2* and **3* being the most common low-activity alleles. Individuals homozygous for either *CYP2C9*2* and *3 or heterozygotes carrying only nonfunctional alleles, clear S-warfarin at a rate that is 90% lower than those with the *CYP2C9*1/*1* genotype. There are differences in the frequencies of these variants, as 8% of Caucasians have been shown to carry at least one copy of the *CYP2C9*2* variants compared to 1% among African Americans, and the allele is rare in Asians. The frequency of *CYP2C9*3* shows a similar distribution, except that it is found in about 3% of Asians [88,89]. Prior pharmacogenomics testing for CYP2C9 variants, coupled with dosage adjustments in Caucasians, reportedly result in a 30% reduction in hospitalizations [90].

In addition to CYP2C9, the warfarin target VKOR is an important determinant of warfarin response because its gene also exhibits genetic polymorphism. The anticoagulant effects of warfarin are exerted through decreased synthesis of vitamin K–dependent clotting factors by inhibiting the *VKORC1* C1 subunit. A mutation (G>A) in the promoter of the *VKCOR1* gene results in individuals presenting with a genotype (AA) needing 50% lower warfarin doses to exert the same effect. Taken together, variation in CYP2C9 (~15%) and VKORC1 (~25%) accounts for nearly 40% of the interindividual variability in warfarin dose requirements among Caucasians [88].

The contribution of each gene differs in different populations as well. However, similar to CYP2D6, the frequency of the homozygous AA, heterozygous AG, and homozygous GG genotype differs in different populations—14%, 47%, and 39% among Caucasians; 82%, 18%, and 0.35% among Asians, and poorly characterized among Africans. This has huge implications for warfarin response in these populations. Other genes, such as CYP4F2 and gamma-glutamyl carboxylase, have also been shown to influence warfarin dose requirements, yet there are no data on their contribution [88,91]. Many algorithms have been developed for warfarin dosing, but very few are applicable to African populations because of lack of information.

In the clinical use of warfarin, based on international normalized ratios (INR) between 2.0 and 3.0, studies show that Asian Americans require low weekly doses of 24 mg/week, followed by Caucasians and Hispanics at 36 and 31 mg/week, respectively. African Americans require the highest dose—an average of 43 mg/week [92]. Therefore, African American ancestry has been used as a predictor of high warfarin dosage [93]. Pharmacogenetic studies based on *CYP2C9* and *VKCORI* are shedding light on these clinical observations of variation in different population groups. In particular, the fact that the *CYP2C9*2* and **3* alleles occur at very low frequencies in African populations [33] means that the FDA-approved pharmacogenetic diagnostic kit for them is of little value to African populations.

N-Acetyl Transferase-2 (NAT-2) in the Disposition of Isoniazid

The variability of isoniazid disposition was one of the first to be characterized and continues to be of great importance to the people of Africa. This is because isoniazid is one of the key components of the cocktail of drugs used to treat tuberculosis. Genetic polymorphism studies led to the discovery of variants that explain rapid acetylator (RA) and slow acetylator (SA) status, which have subsequently been studied in African populations [94]. Pharmacometric modeling of isoniazid clearance in Caucasian SAs and RAs shows that the currently used standard dose can be doubled in RA and halved in SA to achieve the same safe and efficacious levels in blood [95].

In the intention-to-treat (ITT) analysis, isoniazid drug-induced liver injury (INH-DILI) occurred in 78% of SAs in the standard (STD) treatment group, while no SA in the PGx-guided group experienced either INH-DILI or early treatment failure. Among RAs, early treatment failure had a significantly lower incidence rate in the PGx group than in the STD group (15.0% vs. 38%). Thus, the NAT-2 genotype-guided regimen resulted in much lower incidence of unfavorable events, INH-DILI, or early treatment failure than the conventional standard regimen [96]. These studies are relevant to African populations for whom the SA phenotype has been shown to exceed 50%, who carry the heaviest burden of TB infections in the world, and are likely to be treated with NAT-2 drug substrates.

43.5 FUTURE PERSPECTIVES

43.5.1 Epigenetics

Even in well-studied Caucasian populations, after taking into account most of the known genetic variation in DME genes, there has always been a gap between genotype-predicted phenotypes and those determined by *in vivo* probe drugs—specifically, inadequate correlation between genotypes and variation in plasma drug levels. As described by Gomez and Ingelman-Sundberg:

Epigenetics describes the changes that do not involve any alteration in DNA sequence and include covalent modifications of DNA and histones, DNA packaging around nucleosomes, chromatin folding and attachment to the nuclear matrix, and regulatory non-coding RNAs, e.g., microRNAs (miRNAs) [97].

Thus, many early observations of environmental effects can now be explained by epigenetics. There are several processes involved in epigenetics, but the most common are DNA methylation and microRNA regulation. For example, methylation in DME promoter regions is associated with reduced expression of the respective genes, while the presence or absence of microRNA target sequences in the 3′UTR region of DME genes results in differential regulation of the same genes [97,98].

In terms of ongoing research, epigenetics has opened up a huge area of interest for many biological disciplines including pharmacogenetics. In addition to traditional pharmacogenetics, which focuses on variation in exon/introns, a new discipline, pharmacoepigenetics, looks at other mechanism to explain interindividual drug response variations. Many of the genes encoding DMEs, drug transporters, transcription factors, drug targets, and nuclear receptors are regulated at the epigenetic as well as the pharmacogenetic level. Studies are now looking at the role of microRNAs (miRNAs) in terms of variation in their coding sequences (mirSNP) and regulatory target sequences. All of this variation affects plasma drug levels and is important in establishing pharmacogenomics algorithms for drug response. DNA methylation has been demonstrated or predicted in the following clinically relevant CYPs, CYP2C19, CYP2D6, and CYP3A4/5. Many studies have reported DNA methylation in GSTs, especially the glutathione-S-transferase P1 (GSTP1) that has been found to be differentially methylated in different tumors. These observations in DMEs are also being reported in nuclear receptors, which are vital in the transcription control of genes encoding drug transporters and enzymes.

Thus, pharmacogenomics knowledge cannot be complete without incorporating effects associated with epigenetic changes, such as DNA methylation and noncoding regulatory RNAs, including microRNAs. For example, a study by Turrin and coworkers [99] reported an inverse correlation between miR212, 328, and ABCG2 (a transporter) gene expression, indicating that variations in miRNAs that regulate genes for drug-metabolizing enzymes, transporters, and targets may give rise to variations in drug response. An appreciable number of chemical entities that affect the genome's DNA methylation landscape are now being used in clinical practice; however, the challenge here is overcoming their lack of specificity. As yet no studies in African populations have reported on the role of epigenetics in drug response, but there is evidence from functional studies that it does have a role to play.

43.5.2 From Pharmacogenetics to Pharmacogenomics: Where Does Africa Feature?

Current advances in genomics research and in related technologies make it possible to translate research results into potential health benefits through pharmacogenomics, ultimately leading to stratified medicine and personalized therapy. SNP functional data are being curated in the encyclopedia of DNA elements (ENCODE), which will aid in the description of functions for most SNPs in the human genome, and thus further our understanding of drug response. Individualization of drug therapy based on genetic make-up is a desirable paradigm, especially for the treatment of diseases with drugs that (1) exhibit narrow therapeutic windows, (2) may raise toxicity above certain levels, or (3) lower therapeutic failure below certain levels (e.g., low levels of ARVs that may promote virus resistance). Currently, antiretroviral and anticancer drugs are administered at maximally tolerated doses as defined for particular populations. Inevitably, up to 30% of patients are thought to present with ADEs, some of them fatal.

Pharmacogenomics can improve prediction of individuals likely to be at greater risk of drug-induced ADEs [100]. However, clinical data on African populations are lacking. For this reason, in addition to embracing pharmacogenomics, epidemiological research should include the mapping of disease and ADE/ADR determinants.

Based mostly on studies on European populations, several pharmacogenetics prognostic tools have been developed or are projected to be developed by 2020 [101]. Some of them have been recommended by the FDA, including the ampliChip CYP450 by Roche, the DMET chip by Affymetrix, and TheraGuide 5-FU by Myriad Genetics. These tools predict either efficacy or toxicity, and most of them at present remain largely irrelevant to African populations because they are not premised on African-specific genetic variants. A recent study evaluated the ampliChip CYP450 tests designed by Roche and reported poor sensitivity when used in a black South African cohort [74]. In general, African populations are likely to benefit more from denser genetic testing tools because of their small LD blocks compared to other world populations. As can be seen, understanding the pharmacogenomic determinants of drug response and toxicity is critical.

Despite considerable advances in our understanding of human genetic architecture, the genetic underpinning of complex disorders remains elusive. On the basis of our current understanding of neurobiology, numerous candidate genes have been implicated in the etiology of and response to treatment of addiction. Genome-wide association studies

(GWASs) have identified other novel targets. However, study replication is often lacking, which complicates interpretation. The situation is expected to improve as issues such as phenotypic characterization, the apparent "missing heritability," the identification of functional variants, and possible gene–environment (G-E) interactions are resolved.

It has become apparent that genetic diversity and heterogeneity in African populations call for comprehensive genomic, including pharmacogenomic, research in as many African populations as possible. Three initiatives, the Malariagen Project (www.malariagen.net), 1000 Genomes (www.1000genomes.org), and H3Africa (www.h3africa.org), are at the forefront of research on African populations. For example, 1000 Genomes is making available a plethora of data from previously uncharacterized population groups. However, the boldest initiative is the Human Heredity and Health for Africa (H3Africa) project.

43.5.3 Catalyzing Genomics Research in Africa: Human Heredity and Health

Human Heredity and Health for Africa (H3Africa) is an initiative born out of the partnership of the Africa Society for Human Genetics (AfSHG), the Wellcome Trust (United Kingdom), and the National Institutes of Health (NIH). Its goal is to redress the disproportionate lack of genomics research in Africa that developed after WHO identified genomics as one of the critical drivers in improving human health. Africa is faced with the quadruple burden of infectious disease, HIV/AIDS, malaria, and TB, as well as the rapid rise in noncommunicable diseases. According to an H3Africa white paper: "Differences in disease prevalence between African countries, geographic regions within a country, and ethnic groups in the same geographic region are the result of a combination of human genomic and environmental factors." Thus, H3Africa seeks to use recent technology advances to strengthen African research in order to identify genetic variations associated with disease susceptibility, pathogenesis, and treatment response. It is leveraging funding from the Wellcome Trust and NIH to promote African genomics research.

Already, several groups have been identified for genomics research funding (see www.h3africa.org). The inclusion of African populations in this global research agenda is of obvious research and practical importance because African populations harbor the greatest genetic diversity, from which all other world populations originated. Thus, genomic research among Africans is likely to lead to new discoveries and advances that can contribute to improved global health. Pharmacogenomics is an important component of the H3Africa initiative, with the expectation that African populations will benefit considerably from pharmacogenomics findings that are likely to reduce ADEs, which are estimated to cost billions of dollars annually.

With the focus of H3Africa on characterizing diversity within indigenous African populations, it is hoped that the availability of data from genomics studies will tell us how genomic variation affects individual and population differences in response to drugs. Studies of genetic variation patterns in African populations are crucial to our understanding of how genes affect phenotypic variation, including disease predisposition and response to treatment.

43.6 CONCLUSIONS

Many more pharmacogenetic and clinical studies are needed to define and characterize the great genetic diversity of African populations with respect to drug efficacy and safety. Pharmacogenomic tools need to be developed to suit the variations seen in African populations to better define dose regimens.

Although pharmacogenomics profiles of drug response and tolerability have been an important area of research in personalized medicine, a huge effort must be exerted toward clinical validation of the myriad of pharmacogenomics tests that will soon be available. Clinical validity will only become possible once we understand and take into account SNPs, copy number variants, noncoding RNA in gene regulation, gene–gene and gene–environmental interactions, and epigenetic modifications.

For the effective translation of pharmacogenomics research, healthcare practitioners should be familiar with genetic determinants of drug response and willing to catalyze the transition of knowledge from bench to bedside and vice versa. Thus, the research and clinical capacities of African researchers and physicians should be strengthened through pharmacogenomics training so that this knowledge can be integrated into medical practice and eventually lead to personalized medicine.

KEY CONCEPTS

- Population diversity manifests itself through genetic, ethnolinguistic and cultural differences as well as differences arising from environmental and sociopolitical factors such as exposure to pathogens, disease burden, diet, climate, poverty, urban versus rural lifestyle, and so on.
- African populations exhibit greater diversity than do populations from Europe, Asia, and the Americas.
- Measurement of population diversity is complicated and may be based on ethnolinguistic–cultural factors or on genetic factors. One nongenetic measure is a country's ethnic fractionalization (EF), defined as the probability that two individuals selected at random from a country will be from different ethnic groups.

- Population structure and genetic diversity have been determined by various approaches, including rates of microsatellite mutation, heterozygosity, linkage disequilibrium, and F-statistics. All of these have led to a detailed understanding of population clusters in Africa and of the extent of their differences, and to the postulation of their migrations both within the continent and out of Africa to Europe, Asia, and the Americas.
- Genetic variation in African populations has arisen from human interaction with various environmental factors, including disease. Genetic and genomics studies have now identified biomarkers for disease susceptibility and/or resistance that influence the practice of medicine. A degree of variation has been postulated to have derived from human interaction with plant life; this so-called animal–plant warfare has resulted in the evolution of xenobiotic-metabolizing enzymes in humans, which are thought to protect humans from plant-derived chemical insults.
- The discovery of genetic polymorphisms in xenobiotic-metabolizing enzymes gave rise to the study of pharmacogenetics. These enzymes have been demonstrated to be largely responsible for the metabolism and disposition of most therapeutics used in disease prevention and treatment.
- The population diversity demonstrated for genes and noncoding regions is also observed for genes encoding drug-metabolizing enzymes and drug transporters. Over decades of phenotype and genotype studies, many clinically important pharmacogenetic variations have been discovered and characterized, such that there are now a number of FDA-approved pharmacogenetic tests and a number of drug labels now bearing information on pharmacogenetics.
- Pharmacogenetic studies in African populations are on the increase and are already demonstrating continental diversity that has relevance in stratified medicine for both indigenous and diaspora Africans. While generalizations in Caucasian populations on the clinical utility of pharmacogenetic traits and tests are possible, they are likely to be difficult with African populations given Africa's demonstrated continental diversity.
- To date, clinical applications of pharmacogenetics have focused mainly on diseases prevalent in Caucasian and Asian populations. In Africa, these applications must be extended to the treatment of numerous viral, bacterial, protozoal, and helminthic diseases.
- Epigenetic changes may show diversity across populations depending on life experiences, and this diversity may have implications for pharmacogenomics and stratified medicine. Given the diverse cultures, diets, disease prevalence, and environments within the African continent, epigenetic studies are likely to produce interesting results.

ACKNOWLEDGMENTS

Funding for this work was from the European and Developing countries Clinical Trial Partnership (EDCTP) and the International Science Program (ISP) for CM and JAH and from the Medical Research Council (MRC) of South Africa, National Research Foundation (NRF), and University of Cape Town for CD. CM is an EDCTP Senior Clinical Research Fellow. We thank the technical staff of the Laboratory for Genotyping Development at RIKEN Center for Genomic Medicine, for genome-wide SNP genotyping in the unpublished work by Maimbo and colleagues [34]. This work was conducted as a part of the Biobanking and Pharmacogenetics project at AiBST, which was supported by EDCTP (European Union) and IPICS (Sweden).

GLOSSARY

Ethnic fractionalization (EF) the probability that two individuals selected at random from a country will be from different ethnic groups

F-statistics (F_{ST}) the proportion of the total genetic variance contained in a subpopulation (subscript S) relative to the total genetic variance (subscript T). Values can range from 0 to 1. High F_{ST} implies a considerable degree of differentiation among populations. F_{IS} (inbreeding coefficient) is the proportion of the variance in the subpopulation contained in an individual. High F_{IS} implies a high degree of inbreeding

Linkage disequilibrium (LD) the tendency for two alleles to be present on the same chromosome (positive LD) or not to move together (negative LD). As a result, specific alleles at two different loci are found together more or less than expected by chance. The same situation may exist for more than two alleles. LD magnitude is expressed as the delta (D) value and corresponds to the difference between the expected and the observed haplotype frequency. It can have positive or negative values and is decreased by recombination, so it decreases every generation of random mating unless some process opposes the approach to linkage equilibrium. Permanent LD may result from natural selection if some gametic combinations result in higher fitness than other combinations

ACRONYMS AND ABBREVIATIONS

ADE	adverse drug event
ADR	adverse drug reaction
AfSHG	African Society for Human Genetics
EF	ethnic fractionalization
EM	extensive metabolizer
ENCODE	Encyclopedia of DNA Elements
CF	cultural fractionalization
CYP	cytochrome P-450
EFV	efavirenz
FDA	Food and Drug Administration
5-FU	5 fluorouracil
GWAS	genome-wide association study
HAART	highly active antiretroviral therapy
H3Africa	Human Heredity and Health for Africa project
INH-DILI	isoniazid drug induced liver injury
LD	linkage disequilibrium
NAT-2	(v) N-Acetyl Transferase-2
NIH	National Institutes of Health
NRTI	nucleoside reverse transcriptase inhibitor

PM	poor metabolizer
PCA	principal component Analysis
SNP	single-nucleotide polymorphism
TB	tuberculosis
UM	ultrarapid metabolizer
VKORC1	vitamin K 2,3-epoxide reductase C1 subunit

REFERENCES

[1] Cavalli-Sforza LL, Piazza A, Menozzi P. History and geography of human genes. Princeton University Press; 1994.

[2] Campbell MC, Tishkoff SA. The evolution of human genetic and phenotypic variation in Africa. Curr Biol 2010;2010(20):R166–73.

[3] Prugnolle F, Manica A, Balloux F. Geography predicts neutral genetic diversity of human populations. Curr Biol 2005;15(5):R159–60.

[4] Betti L, Balloux F, Amos W, Hanihara T, Manica A. Distance from Africa, not climate, explains within-population phenotypic diversity in humans. Proc Biol Sci 2009;276(1658):809–14.

[5] Tishkof SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, et al. The genetic structure and history of Africans and African Americans. Science 2009;324(5930):1035–44.

[6] Lock M. The art of medicine—the lure of the epigenome. Lancet 2013;381:1896–7.

[7] Labonté B, Suderman M, Maussion G, Navaro L, Yerko V, Mahar I, et al. Genome-wide epigenetic regulation by early-life trauma. Arch Gen Psychiatry 2012;69:722–31.

[8] Meaney MJ. Epigenetics and the biological definition of gene x environment interactions. Child Dev 2010;81:41–79.

[9] Moshe S. The early life social environment and DNA methylation. Epigenetics 2011;6:1–8.

[10] Nieminen TT, Shoman S, Eissa S, Peltomäki P, Abdel-Rahman WM. Distinct genetic and epigenetic signatures of colorectal cancers according to ethnic origen. Cancer Epidemiol Biomarkers Prev 2012;21(202–211):011.

[11] Mobio TA, Anane R, Baudrimont I, Carratú MR, Shier TW, Dano SD, et al. Epigenetic properties of fumonisin B(1): cell cycle arrest and DNA base modification in C6 glioma cells. Toxicol Appl Pharmacol 2000;164(1):91–6.

[12] Webb JLA. Historical epidemiology and infectious disease processes in Africa. J Afr Hist 2013;54:3–10.

[13] African world heritage—a remarkable diversity, UNESCO Publishing; 2013. <http://publishing.unesco.org/details.aspx?&Code_Livre=4959&change=E>.

[14] Cashdan E. Ethnic diversity and its environmental determinants: effects of climate change, pathogens, and habitat diversity. Am Anthropol 2001;103:968–91.

[15] Heine B, Nurse D, editors. African languages: an introduction. Cambridge University Press; 2000.

[16] Fearon JD. Ethnic and cultural diversity by country. J Econ Growth 2003;8:195–222.

[17] Posner D. Measuring ethnic fractionalization in Africa. Am J Polit Sci 2004;48(4):849–63.

[18] Roeder PG. Ethnolinguistic fractionalization (ELF) indices, 1961 and 1985. 2001 <http//:weber.ucsd.edu/~proeder/elf.htm> [24.04.2003].

[19] Alesina A, Devleeschauwer A, Easterly W, Kurlat S, Wacziarg R. Fractionalization. J Econ Growth 2003;8:155–94.

[20] Campbell MC, Tishkoff SA. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet 2008;9:403–33.

[21] Rosa A, Brehm A. African human mtDNA phylogeography at-a-glance. J Anthropol Sci 2011;89:25–58.

[22] UN Department of Economic and Social Affairs, Population Division. See <http//:unstats.un.org/demographic/products/socind/> and <http://esa.un.org/wup2009/unup/index.asp?panel=1>.

[23] WHO Health Data and Statistics, <www.who.int/healthinfo/statistics/en/>.

[24] UN Department of Economic and Social Affairs, Population Division. Population aging and development 2012, Wall chart. <www.un.org/en/development/desa/population/publications/pdf/ageing/2012PopAgeingandDev_WallChart.pdf>.

[25] WHO Global burden of disease, <www.who.int/topics/global_burden_of_disease/en/>.

[26] WHO World Health Statistics. Global health indicators part III, WHO report, 2013 <www.who.int/gho>.

[27] WHO Cause-specific mortality and morbidity: HIV/AIDS, malaria and TB. <http://apps.who.int/gho/data/node.main.27?lang=en>.

[28] Taylor AL, Ziesche S, Yancy C, Carson P, D'agostino Jr R, Ferdinand K, et al. African-American heart failure trial investigators. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. N Engl J Med 2004;351:2049–57.

[29] Exner DV, Dries DL, Domanski MJ, Cohn JN. Lesser response to angiotensin-converting-enzyme inhibitor therapy in blacks as compared with white patients with left ventricular dysfunction. N Engl J Med 2001;344:1351–7.

[30] Zoratti R. A review on ethnic differences in plasma triglycerides and high-density-lipoprotein cholesterol: is the lipid pattern the key factor for the low coronary heart disease rate in people of African origin? Eur J Epidemiol 1998;14(1):9–21.

[31] Gaines K, Burke G. Ethnic differences in stroke: black-white differences in the United States population. SECORDS investigators. Southeastern consortium on racial differences in stroke. Neuroepidemiology 1995;14(5):209–39.

[32] Scheinfeldt LB, Soi S, Tishkoff SA. Working towards a synthesis of archeological, linguistic, and genetic data for inferring African population history. PNAS 2010;107:8931–8.

[33] Matimba A, Oluka MN, Ebeshi BU, Sayi J, Bolaji OO, Guantai AN, et al. Establishment of a biobank and pharmacogenetics database of African populations. Eur J Hum Genet 2008;16:780–3.

[34] Maimbo M, Mushiroda T, Takahashi A, Kubo M, Sayi J, Bolaji OO, et al. A comprehensive genome-wide study of the genetic structure of individuals from 10 African ethnic groups, manuscript in preparation.

[35] Bryc K, Auton A, Nelson MR, Oksenberg JR, Hauser SL, Williams S, et al. Genome-wide patterns of population structure and admixture in West Africans and African Americans. Proc Natl Acad Sci USA 2010;12:786–91.

[36] Fage JD. A history of Africa. In: West Africa during the slave trade era. Unwin Hyman; 1990. p. 266 [Chapter 11].

[37] Tishkoff SA, Gonder MK, Henn BM, Mortensen H, Knight A, Gignoux C, et al. History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. Mol Biol Evol 2007;24:2180–95.

[38] Ambrose S. Chronology of the Later Stone Age and food production in East Africa. J Arch Sci 1998;25:377–91.

[39] Gifford-Gonzalez D. African archeology. Blackwell 2005:187–224.

[40] Hansen NT, Brunak S, Altman RB. Generating genome-scale candidate gene lists for pharmacogenomics. Clin Pharmacol Ther 2009;86(2):183–9.

[41] Kitzmiller JP, Groen DK, Phelps MA, Sadee W. Pharmacogenomic testing: relevance in medical practice: why drugs work in some patients but not in others. Cleve Clin J Med 2011;78(4):243–57.

[42] Hernandez-Boussard T, Woon M, Klein TE, Altman RB. Integrating large-scale genotype and phenotype data. OMICS 2006;10(4):545–54.

[43] Sun HY, Hung CC, Lin PH, Chang SF, Yang CY, Chang SY, et al. Incidence of abacavir hypersensitivity and its relationship with HLA-B*5701 in HIV-infected patients in Taiwan. J Antimicrob Chemother 2007;60(3):599–604.

[44] Kudzi W, Dodoo AN, Mills JJ. Characterisation of CYP2C8, CYP2C9 and CYP2C19 polymorphisms in a Ghanaian population. BMC Med Genet 2009;10:124.

[45] Kudzi W, Dodoo AN, Mills JJ. Genetic polymorphisms in MDR1, CYP3A4 and CYP3A5 genes in a Ghanaian population: a plausible explanation for altered metabolism of ivermectin in humans? BMC Med Genet 2010;11:111.

[46] Habtewold A, Amogne W, Makonnen E, Yimer G, Riedel KD, Ueda N, et al. Long-term effect of efavirenz autoinduction on plasma/peripheral blood mononuclear cell drug exposure and CD4 count is influenced by UGT2B7 and CYP2B6 genotypes among HIV patients. J Antimicrob Chemother 2011;66(10):2350–61.

[47] Ngaimisi E, Habtewold A, Minzi O, Makonnen E, Mugusi S, Amogne W, et al. Importance of ethnicity, CYP2B6 and ABCB1 genotype for efavirenz pharmacokinetics and treatment outcomes: a parallel-group prospective cohort study in two sub-Saharan Africa Populations. PLoS ONE 2013;8(7).

[48] Swart M, Ren Y, Smith P, Dandara C. ABCB1 4036A>G and 1236C>T Polymorphisms affect plasma efavirenz levels in South African HIV/AIDS patients. Front Genet 2012;3:236.

[49] Swart M, Skelton M, Ren Y, Smith P, Takuva S, Dandara C. High predictive value of CYP2B6 SNPs for steady-state plasma efavirenz levels in South African HIV/AIDS patients. Pharmacogenet Genomics 2013;23(8):415–27.

[50] Swart M, Whitehorn H, Ren Y, Smith P, Ramesar RS, Dandara C. PXR and CAR single nucleotide polymorphisms influence plasma efavirenz levels in South African HIV/AIDS patients. BMC Med Genet 2012;13:112.

[51] Yimer G, Amogne W, Habtewold A, Makonnen E, Ueda N, Suda A, et al. High plasma efavirenz level and CYP2B6*6 are associated with efavirenz-based HAART-induced liver injury in the treatment of naïve HIV patients from Ethiopia: a prospective cohort study. Pharmacogenomics J 2012;12(6):499–506.

[52] Kampira E, Kumwenda J, van Oosterhout JJ, Dandara C. Mitochondrial DNA subhaplo groups L0a2 and L2a modify susceptibility to peripheral neuropathy in Malawian adults on stavudine containing highly active antiretroviral therapy. J Acquir Immune Defic Syndr 2013;63:647–52.

[53] Mukonzo JK, Okwera A, Nakasujja N, Luzze H, Sebuwufu D, Ogwal-Okeng J, et al. Influence of efavirenz pharmacokinetics and pharmacogenetics on neuropsychological disorders in Ugandan HIV-positive patients with or without tuberculosis: a prospective cohort study. BMC Infect Dis 2013;13:261.

[54] Rendic S, Guengerich FP. Update information on drug metabolism systems–2009, part II: summary of information on the effects of diseases and environmental factors on human cytochrome P450 (CYP) enzymes and transporters. Curr Drug Metab 2010;11(1):4–84.

[55] Lynch T, Price A. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. Am Fam Physician 2007;76(3):391–6.

[56] Bains RK, Kovacevic M, Plaster CA, Tarekegn A, Bekele E, Bradman NN, et al. Molecular diversity and population structure at the cytochrome P450 3A5 gene in Africa. BMC Genet 2013;14:34.

[57] Swen JJ, Nijenhuis M, DeBoer A, Grandia L, Maitland-van der Zee AH, Mulder H, et al. Pharmacogenetics: from bench to byte—an update of guidelines. Clin Pharmacol Ther 2011;89(5):662–73.

[58] Alessandrini M, Asfaha S, Dodgen TM, Warnich L, Pepper MS. Cytochrome P450 pharmacogenetics in African populations. Drug Metab Rev 2013;45(2):253–75.

[59] Masimirembwa CM, Hasler JA. Genetic polymorphism of drug metabolising enzymes in African populations: implications for the use of neuroleptics and antidepressants. Brain Res Bull 1997;44(5):561–71.

[60] Aklillu E, Dandara C, Bertilsson L, Masimirembwa C. Pharmacogenetics of cytochrome P450s in African populations: clinical and molecular evolutionary implications. Chapter category: pharmacogenomics. In: Suarez-Kurtz Guilherme, editor. Pharmacogenomics in admixed populations. Landes Bioscience; 1990. p. 99 [Chapter 7].

[61] Masimirembwa CM, Johansson I, Hasler JA, Ingelman-Sundberg M. Genetic polymorphism of cytochrome P450 CYP2D6 in Zimbabwean population. Pharmacogenetics 1993;3(6):275–80.

[62] Masimirembwa C, Bertilsson L, Johansson I, Hasler JA, Ingelman-Sundberg M. Phenotyping and genotyping of S-mephenytoin hydroxylase (cytochrome P450 2C19) in a Shona population of Zimbabwe. Clin Pharmacol Ther 1995;57(6):656–61.

[63] Masimirembwa C, Persson I, Bertilsson L, Hasler J, Ingelman-Sundberg M. A novel mutant variant of the CYP2D6 gene (CYP2D6*17) common in a black African population: association with diminished debrisoquine hydroxylase activity. Br J Clin Pharmacol 1996;42(6):713–9.

[64] Masimirembwa C, Hasler J, Bertilssons L, Johansson I, Ekberg O, Ingelman-Sundberg M. Phenotype and genotype analysis of debrisoquine hydroxylase (CYP2D6) in a black Zimbabwean population. Reduced enzyme activity and evaluation of metabolic correlation of CYP2D6 probe drugs. Eur J Clin Pharmacol 1996;51(2):117–22.

[65] Dandara C, Masimirembwa CM, Magimba A, Sayi J, Kaaya S, Sommers DK, et al. Genetic polymorphism of CYP2D6 and CYP2C19 in East- and Southern African populations including psychiatric patients. Eur J Clin Pharmacol 2001;57(1):11–17.

[66] Griese EU, Asante-Poku S, Ofori-Adjei D, Mikus G, Eichelbaum M. Analysis of the CYP2D6 gene mutations and their consequences for enzyme function in a West African population. Pharmacogenetics 1999;9(6):715–23.

[67] Bathum L, Skjelbo E, Mutabingwa TK, Madsen H, Hørder M, Brøsen K. Phenotypes and genotypes for CYP2D6 and CYP2C19 in a black Tanzanian population. Br J Clin Pharmacol 1999;48(3):395–401.

[68] Herrlin K, Massele AY, Jande M, Alm C, Tybring G, Abdi YA, et al. Bantu Tanzanians have a decreased capacity to metabolize omeprazole and mephenytoin in relation to their CYP2C19 genotype. Clin Pharmacol Ther 1998;64(4):391–401.

[69] Aklillu E, Herrlin K, Gustafsson LL, Bertilsson L, Ingelman-Sundberg M. Evidence for environmental influence on CYP2D6-catalysed debrisoquine hydroxylation as demonstrated by phenotyping and genotyping of Ethiopians living in Ethiopia or in Sweden. Pharmacogenetics 2002;12(5):375–83.

[70] Simooya OO, Njunju E, Hodjegan AR, Lennard MS, Tucker GT. Debrisoquine and metoprolol oxidation in Zambians: a population study. Pharmacogenetics 1993;3(4):205–8.

[71] Simooya OO, Sijumbil G, Lennard MS, Tucker GT. Halofantrine and chloroquine inhibit CYP2D6 activity in healthy Zambians. Br J Clin Pharmacol 1998;45(3):315–7.

[72] Gaedigk A, Coetsee C. The CYP2D6 gene locus in South African Coloureds: unique allele distributions, novel alleles and gene arrangements. Eur J Clin Pharmacol 2008;64(5):465–75.

[73] Wright GE, Niehaus DJ, Drögemöller BI, Koen L, Gaedigk A, Warnich L. Elucidation of CYP2D6 genetic diversity in a unique African population: implications for the future application of pharmacogenetics in the Xhosa population. Ann Hum Genet 2010;74(4):340–50.

[74] Dodgen TM, Hochfeld WE, Fickl H, Asfaha SM, Durandt C, Rheeder P, et al. Introduction of the AmpliChip CYP450 test to a South African cohort: a platform comparative prospective cohort study. BMC Med Genet 2013;14:20.

[75] Wennerholm A, Dandara C, Sayi J, Svensson JO, Abdi YA, Ingelman-Sundberg M, et al. The African-specific CYP2D6*17 allele encodes an enzyme with changed substrate specificity. Clin Pharmacol Ther 2002;71(1):77–88.

[76] Bapiro TE, Hasler JA, Ridderström M, Masimirembwa CM. The molecular and enzyme kinetic basis for the diminished activity of the cytochrome P450 2D6.17 (CYP2D6.17) variant. Potential implications for CYP2D6 phenotyping studies and the clinical use of CYP2D6 substrate drugs in some African populations. Biochem Pharmacol 2002;64:1387–98.

[77] Thompson EE, Kuttab-Boulos H, Witonsky D, Yang L, Roe BA, Di Rienzo A. CYP3A variation and the evolution of salt-sensitivity variants. Am J Hum Genet 2004;75(6):1059–69.

[78] Byakika-Kibwika P, Lamorde M, Mayito J, Nabukeera L, Namakula R, Mayanja-Kizza H, et al. Significant pharmacokinetic interactions between artemether/lumefantrine and efavirenz or nevirapine in HIV-infected Ugandan adults. J Antimicrob Chemother 2012;67(9):2213–21.

[79] Nemaura T, Nhachi C, Masimirembwa C. Impact of gender, weight and CYP2B6 genotype on efavirenz exposure in patients on HIV/AIDS and TB treatment: implications for individualising therapy. Afr J Pharm Pharmacol 2012;6(29):2188–93.

[80] di Iulio J, Fayet A, Arab-Alameddine M, Rotger M, Lubomirov R, Cavassini M, et al. Swiss HIV cohort study. *In vivo* analysis of efavirenz metabolism in individuals with impaired CYP2A6 function. Pharmacogenet Genomics 2009;19(4):300–9.

[81] Gounden V, van Niekerk C, Snyman T, George JA. Presence of the CYP2B6 516G> T polymorphism, increased plasma efavirenz concentrations and early neuropsychiatric side effects in South African HIV-infected patients. AIDS Res Ther 2010;7:32.

[82] Nyakutira C, Röshammar D, Chigutsa E, Chonzi P, Ashton M, Nhachi C, et al. High prevalence of the CYP2B6 516G–>T(*6) variant and effect on the population pharmacokinetics of efavirenz in HIV/AIDS outpatients in Zimbabwe. Eur J Clin Pharmacol 2008;64(4):357–65.

[83] Maimbo M, Kiyotani K, Mushiroda T, Masimirembwa C. Y N: CYP2B6 genotype is a strong predictor of systemic exposure 5 to efavirenz in HIV-infected Zimbabweans. Eur J Clin Pharmacol 2012;68:267–71.

[84] Stingl JC, Brockmöller J, Viviani R. Genetic variability of drug-metabolizing enzymes: the dual impact on psychiatric therapy and regulation of brain function. Mol Psychiatry 2013;18(3):273–87.

[85] Hicks JK, Swen JJ, Thorn CF, Sangkuhl K, Kharasch ED, Ellingrod VL, et al. Clinical pharmacogenetics implementation consortium guideline for CYP2D6 and CYP2C19 genotypes and dosing of tricyclic antidepressants. Clinical pharmacogenetics implementation consortium. Clin Pharmacol Ther 2013;93(5):402–8.

[86] Wennerholm A, Johansson I, Hidestrand M, Bertilsson L, Gustafsson LL. Characterization of the CYP2D6*29 allele commonly present in a black Tanzanian population causing reduced catalytic activity. Pharmacogenetics 2001;11(5):417–27.

[87] Kirchheiner J, Brockmöller J. Clinical consequences of cytochrome P450 2C9 polymorphisms. Clin Pharmacol Ther 2005;77(1):1–16.

[88] Yin T, Miyata T. Warfarin dose and the pharmacogenomics of CYP2C9 and VKORC1—rationale and perspectives. Thromb Res 2007;120(1):1–10.

[89] Speed WC, Kang SP, Tuck DP, Harris LN, Kidd KK. Global variation in CYP2C8-CYP2C9 functional haplotypes. Pharmacogenomics J 2009;9(4):283–90.

[90] Aithal GP, Day CP, Kesteven PJ, Daly AK. Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet 1999;353(9154):717–9.

[91] Borgiani P, Ciccacci C, Forte V, Sirianni E, Novelli L, Bramanti P, et al. CYP4F2 genetic variant (rs2108622) significantly contributes to warfarin dosing variability in the Italian population. Pharmacogenomics 2009;10(2):261–6.

[92] Dang MT, Hambleton J, Kayser SR. The influence of ethnicity on warfarin dosage requirement. Ann Pharmacother 2005;39:1008–12.

[93] Absher RK, Moore ME, Parker MH. Patient-specific factors predictive of warfarin dosage requirements. Ann Pharmacother 2002;36:1512–7.

[94] Dandara C, Masimirembwa CM, Magimba A, Kaaya S, Sayi J, Sommers DK, et al. Arylamine N-acetyltransferase (NAT2) genotypes in Africans: the identification of a new allele with nucleotide changes 481C>T and 590G>A. Pharmacogenetics 2003;13(1):55–8.

[95] Kinzig-Schippers M, Tomalik-Scharte D, Jetter A, Scheidel B, Jakob V, Rodamer M, et al. Should we use N-acetyltransferase type 2 genotyping to personalize isoniazid doses? Antimicrob Agents Chemother 2005;49(5):1733–8.

[96] Azuma J, Ohno M, Kubota R, Yokota S, Nagai T, Tsuyuguchi K, et al. Pharmacogenetics-based tuberculosis therapy research group. NAT2 genotype guided regimen reduces isoniazid-induced liver injury and early treatment failure in the 6-month four-drug standard treatment of tuberculosis: a randomized controlled trial for pharmacogenetics-based therapy. Eur J Clin Pharmacol 2013;69(5):1091–101.

[97] Gomez A, Ingelman-Sundberg M. Pharmacoepigenetics: its role in interindividual differences in drug response. Clin pharmacol Therapeut 2009;85(4):426–30.

[98] Baer-Dubowska W, Majchrzak-Celińska A, Cichocki M. Pharmocoepigenetics: a new approach to predicting individual drug responses and targeting new drugs. Pharmacol Rep 2011;63(2):293–304.

[99] Turrini E, Haenisch S, Laechelt S, Diewock T, Bruhn O, Cascorbi I. MicroRNA profiling in K-562 cells under imatinib treatment: influence of miR-212 and miR-328 on ABCG2 expression. Pharmacogenet Genomics 2012;22(3):198–205.

[100] Wang L, McLeod HL, Weinshilboum RM. Genomic medicine: genomics and drug response. N Engl J Med 2011;364:1144–53.

[101] Huang RS, Dolan ME. Approaches to the discovery of pharmacogenomic markers in oncology: 2000–2010–2020. Pharmacogenomics 2010;11(4):471–4.

Chapter 44

Pharmacogenomics in China

Wei Zhang, Hong-Hao Zhou, Zhao-Qian Liu, Xiao-Ping Chen, and Yi-Jing He
Pharmacogenetics Research Institute, Institute of Clinical Pharmacology, Central South University, PR China

Chapter Outline

44.1 Overview 999
44.2 Pharmacogenomics of Drug-Metabolizing Enzymes in Chinese Populations 999
  44.2.1 CYP2C19 999
  44.2.2 CYP2D6 1001
  44.2.3 CYP3A 1001
  44.2.4 Histamine N-Methyltransferase 1001
44.3 Pharmacogenomics of Drug Transporters in Chinese Populations 1002
  44.3.1 ABCC1 1002
  44.3.2 ABCB1 and ABCC2 1003
  44.3.3 ABCG2 1004
  44.3.4 SLCO1B1 1005
  44.3.5 SLC30A8 1006
44.4 Pharmacogenomics of Drug Receptors in Chinese Populations 1006
  44.4.1 β1-Adrenergic Receptor 1006
  44.4.2 β3-Adrenergic rReceptor 1007
  44.4.3 PPAR-γ2 1008
  44.4.4 GABA Receptor 1008
44.5 Pharmacogenomics of ION Channels in Chinese Populations 1009
  44.5.1 KCNH2 1009
  44.5.2 KCNQ1 1010
  44.5.3 KCNJ11 1010
  44.5.4 SCN1A 1010
44.6 Environmental Factors in Drug-Metabolizing Enzyme and Transporter Activity 1011
  44.6.1 CYP2C19-Mediated Herb–Drug Interaction 1011
  44.6.2 OATP1B1-Mediated Herb–Drug Interaction 1011
44.7 Translational Approaches to Pharmacogenomics in China 1012
44.8 Conclusion and Future Perspectives 1012
Acknowledgments 1012
References 1012

44.1 OVERVIEW

Pharmacogenetic and pharmacogenomic studies have established the importance of polymorphic drug-metabolizing enzymes, receptors, transporters, and ion channels in the differential response of patients to drugs. Among these target genes, cytochrome P450 has become the most important subject of extensive studies concerning individual variation in drug metabolism and disposition. The association between decreased drug clearance and the activity of several kinds of CYP enzymes, the inherited nature of the functional deficiency in drug-related genes, and their frequencies and clinical importance were evaluated extensively [1]. During the past ten years, the genetic polymorphisms of transporters and receptors have been increasingly studied at the protein and gene level. It has gradually been demonstrated that genetic differences in ion channels reflect differences in the distribution of polymorphic traits, such as disease susceptibility and drug efficacy or tolerability, which has become a rising field of study on pharmacogenetics and pharmacogenomics research.

In this chapter, we mainly summarize studies with Chinese populations on the individual variations in drug disposition and response mediated by drug-metabolizing enzymes, drug receptors, drug transporters, and ion channels, as well as the role of genotype on herb–drug interactions and translational medicine.

44.2 PHARMACOGENOMICS OF DRUG-METABOLIZING ENZYMES IN CHINESE POPULATIONS

44.2.1 CYP2C19

The genetic polymorphism of CYP2C19 has been elucidated in several different studies of the Chinese population

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00044-X
© 2014 Elsevier Inc. All rights reserved.

TABLE 44.1 CYP2C19 Allele Frequencies in Chinese Subjects

	Han (n = 202)	Dong (n = 148)	Dai (n = 386)	Bai (n = 404)
*CYP2C19*1*	0.559 (113)	0.541 (80)	0.666* (257)	0.688* (278)
*CYP2C19*2*	0.336 (74)	0.399 (59)	0.303 (117)	0.257* (104)
*CYP2C19*3*	0.074 (15)	0.060 (9)	0.031* (12)	0.052 (21)
*CYP2C19*5*	–	–	–	0.0025 (1)

since the 1990s. The two defective alleles were found together, first in poor metabolizers (PMs) of the Chinese Dong nationality, and accounted for 100% of the PM alleles. CYP2C19*2 and CYP2C19*3 represent 86.8% and 13.2% of the mutant alleles in this ethnic group, respectively. Similar results were then defined in the Chinese Han majority (83.2% and l6.8%, respectively) and Chinese Bai minority (82.5% and l6.7%, respectively) [2].

In a group of Han subjects (n = 101), 19.8% were classified as PMs phenotypically and 100% of these phenotypes could be explained by CYP2C19*2 and CYP2C19*3. Several other studies carried out by Shimada et al. and Qin et al. confirmed that CYP2C19*2 accounts for most mutant CYP2C19 alleles in Han subjects, and that CYP2C19*3 can partially explain these mutant alleles in Chinese Bai subjects; 27 of 202 (13.4%) were classified as PMs phenotypically, and only one appeared to be an outlier. The outlier was finally found to be a heterozygote with a CYP2C19*2 and a new mutant allele consisting of a C→T mutant allele at bp 1297 in exon 9. This mutation, which is designated as CYP2C19*5, results in the substitution of Arg433→Trp433 in the heme-binding region and may produce an inactive protein [3].

It has also been determined that CYP2C19*2 and CYP2C19*3 account for all mutant alleles in the Chinese Dai minority (90.7% and 9.3%, respectively). It was noticed that other rare mutations of CYP2C19 found in Caucasian populations—including CYP2C19*6, CYP2C19*7, and CYP2C19*8—were not detected in the Chinese population. The CYP2C19 allele frequencies in these ethnic groups are presented in Table 44.1, in which significant differences in frequencies are clearly shown.

It has been found that there is a large ethnic difference in the amount of CYP2C19 protein in microsomes. The amount of CYP2C19 in microsomes from Japanese and Caucasians is 0.8% and 1.4% of total CYP450, respectively. Moreover, the amount of CYP2C19 in microsomes is highest in individuals who are homozygous for the normal gene (CYP2C19*1/*1), and lowest in those homozygous for the mutation gene (CYP2C19*2/*2 or CYP2C19*3/*3) compared with heterozygous (CYP2C19*1/*2 and CYP2C19*l/*3). Studies that investigated the relationship of gene dose and the ratio of S- and R-mephenytoin in Chinese Han, Dong, Dai, and Bai subjects showed that heterozygous extensive metabolizers (CYP2C19*1/*2 and CYP2C19*1/*3) had significantly higher S/R ratios compared with homozygous wild-type extensive metabolizers (CYP2C19*1/*1). This clearly indicated that gene dosage has an affect on mephenytoin metabolism.

The gene dosage effect of CYP2C19 could influence the metabolism of its substrates. Diazepam is one of the most commonly prescribed sedative drugs for the treatment of anxiety, convulsions, and muscle spasms. N-demethylation is the major metabolic pathway of diazepam in vivo at therapeutic doses. There is evidence that in white and Korean populations the metabolism of both diazepam and its N-demethylated metabolite, desmethyldiazepam, cosegregates with the S-mephenytoin hydroxylation polymorphism; however, the data from studies of the Chinese population conflict with the findings from studies of white and Korean populations. It has been suggested that this discrepancy might be related to the proportion of heterozygotes in Chinese versus Caucasian extensive metabolizers (EMs).

Our studies have shown that the presence of mutations of the CYP2C19 gene cosegregates with the impaired metabolism of diazepam among Chinese subjects in a manner suggesting a gene–dosage effect. A significant difference in the half-lives (t1/2) existed between the heterozygous EMs (CYP2C19*1/*2) and homozygous EMs (CYP2C19*1*1). The slowest mean clearance of diazepam was seen in the CYP2C19*2/*2 subjects, and the fastest in the CYP2C19*1/*1 subjects, with the CYP2C19*1/CYP2C19*2 heterozygotes having an intermediate value. Figure 44.1 shows the demographic characteristics and pharmacokinetic parameters of oral diazepam in healthy Han Chinese subjects with different CYP2C19 genotypes. There were significant intergenotypic differences in the plasma half-lives of diazepam and its active metabolite demethyldiazepam between the different genotypic groups, with the subjects homozygous for the allele m1 (CYP2C19*2/*2) having the longest t1/2 and the wild-type (WT) homozygotes (CYP2C19*1/*1) having the shortest t1/2 for both compounds.

As expected, the slowest mean clearance of diazepam existed in the subjects with CYP2C19*2/*2, and the fastest was observed in the WT/WT subjects. As hypothesized, the heterozygotes (CYP2C19*1/*2) had intermediate values of systemic clearance and elimination half-life. This study was the first definitive evidence that the gene–dosage of CYP2C19 can markedly affect the metabolism and disposition of diazepam and desmethyldiazepam in humans [4].

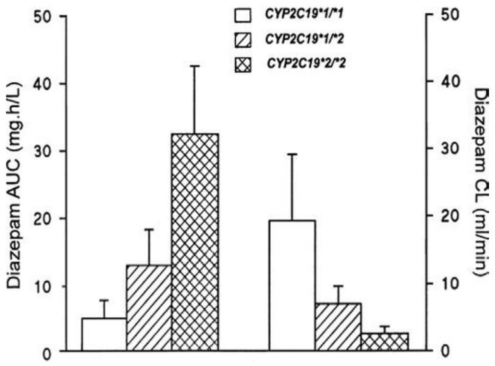


FIGURE 44.1 Effect of CYP2C19 gene dosage on diazepam pharmacokinetic parameters.

44.2.2 CYP2D6

Chinese individuals have greater sensitivity than Caucasians to the effects of propranolol on heart rate and blood pressure. Decreased protein binding and increased sensitivity to plasma renin activity suppression in Chinese can only partially explain their increased sensitivity to propranolol. CYP2D6 is one of the most important drug-metabolizing enzymes in the human liver, which is mainly responsible for propranolol clearance. About 71% of CYP2D6 alleles code for normal enzyme in Caucasians, while 26% is a nonfunctional enzyme mainly prescribed as CYP2D6*4 (15.6%), CYP2D6*5 (6.9%), and CYP2D6*3 (1.6%). Except for CYP2D6*5, which exhibits a similar frequency in Caucasians, Japanese, and Chinese, the other functionally defective alleles of CYP2D6 are very rare in Chinese. However, Chinese metabolize CYP2D6-mediated drugs more slowly than Caucasians, which may due predominantly to high frequencies of variants of 2D6*10—a reduced function allele.

The frequency of mutant allele CYP2D6*10 that encodes a reduced activity of the enzyme is even higher than that of CYP2D6*1 (50% vs. 26.9%), resulting in a higher prevalence of this allelic variant among Chinese patients. Thus, the average activity value of CYP2D6 is lower in Chinese than in other racial populations whose frequencies of the CYP2D6*10 allele are remarkably lower. Therefore, Chinese subjects exhibit a relatively lower metabolizing activity of CYP2D6 on average and may be more sensitive than Caucasians to the effect of CYP2D6 substrates (e.g., propranolol, metoprolol, timolol).

44.2.3 CYP3A

Human cytochrome P450 3A evolved to catalyze the metabolism of numerous common therapy drugs and endogenous molecules. Members of the CYP3A family make up the majority of those expressed in human liver and intestine, and there are marked interindividual differences in their protein expression and activity. The activity of the CYP3A enzyme in the Chinese is highly variable, exceeding 14-fold, and contributes greatly to variation in oral bioavailability and systemic clearance of CYP3A substrates. The genetic factors play an important role in the interindividual variability in CYP3A activity.

Detection of CYP3A5 and CYP3A4 variant alleles and knowledge about their allelic frequency in specific ethnic groups are important in determining individualized drug dosing and improved therapeutics. In a group of 302 unrelated Chinese healthy volunteers, the frequency of the CYP3A5*3 and CYP3A4*18 variant allele were 77.8% and 1%, respectively [5]. The CYP3A4*18 and CYP3A5*3 allelic frequencies are similar to those reported previously for Chinese residents in Taiwan. The frequency of the CYP3A5*3 allele in that population is similar to that in the Japanese but lower than Caucasians. Meanwhile, approximate 62% of the Chinese population carrying the CYP3A5*3/*3 genotype may appear not to express CYP3A5 protein and exhibit a lower total expression of CYP3A.

The calcineurin inhibitor, cyclosporine, is widely used to prevent allograft rejection after solid organ transplantation. It has a narrow therapeutic index and shows considerable interindividual differences in its pharmacokinetics. Interindividual differences in the activity and expression of the metabolizing enzymes cytochrome P450 (CYP) 3A4 and 3A5 contribute considerably to cyclosporine pharmacokinetics. The median cyclosporine dose-adjusted (C0) in CYP3A5*1/*1 genotype subjects was 14.8 ng/mL per mg/kg; in CYP3A5*1/*3 patients ($n = 34$), it was 23.7 ng/mL per mg/kg; for CYP3A5*3/*3 patients ($n = 66$), it was 26.4 ng/mL per mg/kg ($p = 0.012$). Accordingly, cyclosporine dose-adjusted C0 was larger in Chinese CYP3A5 nonexpressors than expressors in the first week after renal transplantation.

The present study shows that genetic polymorphism in CYP3A5 is responsible for the large interindividual variability of cyclosporine pharmacokinetics during the early phase after renal transplantation in Chinese patients. Those with the CYP3A5*3 variant genotype require a low dose of cyclosporine to reach target levels compared with those with the CYP3A5*1 allele.

44.2.4 Histamine N-Methyltransferase

Histamine N-methyltransferase (HNMT) plays an important role in the metabolism of histamine—a biogenic amine that has many physiologic and pathological roles in human tissues. A total of 11 single nucleotide polymorphisms (SNPs) were identified in 192 unrelated healthy Chinese adults, among which six SNPs had variant allele frequencies greater than 5% (Figure 44.2). Of the six common SNPs, three (21637T>C, 2463T>C, and 2411 C>T) were located in 5′-FR, one (314C>T) in coding exons, and two (939A>G and 1097A>T) in the 3′-untranslated region (3′-UTR). Most of these common SNPs were in linkage disequilibrium.

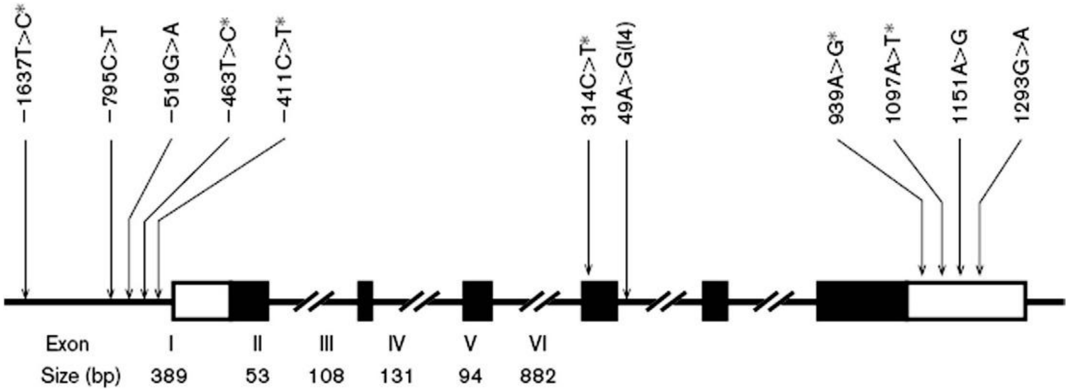


FIGURE 44.2 **Human HNMT structure and polymorphisms in Chinese subjects.** Exons are represented as rectangles: black rectangles—encoding; white—UTR sequences. SNP positions are indicated by arrows; common SNP, by an asterisk. The polymerase chain reaction product sizes corresponding to segments of HNMT are also shown.

Genotype–phenotype correlation analyses were performed for those common SNPs in 5′-FR and 3′-UTR. In males, no significant association was found between HNMT activity and these noncoding SNPs. However, in females, the 21637T>C or 2463T>C tended to be associated with decreased HNMT activity, whereas the 939A>G or 1097A>T appeared to be correlated with increased enzymatic activity. HNMT polymorphisms differ considerably between Chinese and American Caucasians.

The common SNPs in 5′-FR (21637T>C and 2463T>C) and 3′-UTR (939A>G and 1097A>T) might conditionally regulate the activity of HNMT, or it might be genetically linked to unknown mutation(s) underlying the HNMT phenotypic variance. The C314T missense mutation (Thr105Ile) in the HNMT gene has been identified as representing a common functional polymorphism in Caucasians, whereas an A595G (Ile199Val) variant has been reported in one HNMT cDNA from a Japanese subject. The point mutations C314T and A595G within HNMT were both detected in 352 unrelated Chinese Han subjects. None of these subjects contained the A595G mutation, whereas 40 (11.6%) heterozygotes and 1 (0.3%) homozygote for the variant T314 allele were detected.

The frequency of the variant T314 allele in this Chinese population was 0.060 (95% CI: 0.042–0.078), no different from Japanese but significantly lower than American Caucasians. The C314T mutation represents a common functional genetic polymorphism in the Chinese Han population with a variant T314 allele frequency similar to Japanese but lower than American Caucasians, whereas the A595G mutation does not appear to exist in this population [6,7].

We investigated the polymorphisms in the HNMT gene in Chinese to assess their relationship with HNMT activity. HNMT polymorphisms were screened in 192 unrelated subjects by direct sequencing with purified polymerase chain reaction products comprising all six exons, plus splice junctions, as well as approximately 2 kb of the 5-flanking region (5-FR). Erythrocyte HNMT activity was measured by radiochemical microassay. A total of 11 single nucleotide polymorphisms (SNPs) were identified, among which six SNPs had variant allele frequencies greater than 5%. Of the six common SNPs, three (1637T>C, −463T>C, and −411C>T) were located in 5′-FR, one (314C>T) in coding exons, and two (939A>G and 1097A>T) in the 3′-untranslated region (3′-UTR). In females, the −1637T>C or −463T>C tended to be associated with decreased HNMT activity, whereas the 939A>G or 1097A>T appeared to be correlated with increased enzymatic activity. The common SNPs in 5′-FR (−1637T>C and −463T>C) and 3′-UTR (939A>G and 1097A>T) of HNMT might conditionally regulate the activity of HNMT in a Chinese population with a gender difference.

44.3 PHARMACOGENOMICS OF DRUG TRANSPORTERS IN CHINESE POPULATIONS

44.3.1 ABCC1

Multidrug resistance-associated protein 1 (MRP1/ABCC1) has been shown to cause multidrug resistance in model cell lines and in clinical settings. We explored the distribution frequencies of four common single nucleotide polymorphisms of MRP1/ABCC1—Cys43Ser (128G>C), Thr73Ile (218C>T), Arg723Gln (2168G>A), and Arg1058Gln (3173G>A)—in a mainland Chinese population and investigated whether these SNPs affect the expression and function of MRP1/ABCC1. The four candidate SNPs were recreated by site-directed mutagenesis and tested for their effect on MRP1/ABCC1 expression and multidrug resistance (MDR) function in stable transfected HEK293 and CHO-K1 cell lines. Real-time PCR, Western blot, and confocal microscopy were used to determine mRNA, protein expression, and protein trafficking. Last, the effect of mutations on MRP1/ABCC1-mediate drug resistance was determined using a methyl thiazolyl tetrazolium assay.

We found the allelic frequencies of Cys43Ser (128G>C), Thr73Ile (218C>T), Arg723Gln (2168G>A), and Arg1058Gln (3173G>A) in mainland Chinese were 0.5, 1.4, 5.8, and 0.5%, respectively. Among them, the Arg723Gln mutation significantly reduced MRP1/ABCC1-mediated resistance to daunorubicin, doxorubicin, etoposide, vinblastine, and vincristine. The Thr73Ile mutation reduced resistance to methotrexate and etoposide, whereas the Arg1058Gln mutation increased the response of two anthracycline drugs and etoposide in HEK293 and CHO-K1 cells as well as vinblastine and methotrexate in CHO-K1 cells (Figures 44.3 and 44.4) [8].

We subsequently investigated the clinical application of this polymorphism with lung cancer susceptibility and chemotherapy response in a Chinese population. A total of 77 lung cancer patients and 71 cancer-free controls were enrolled in the study. Individuals carrying the 723Gln (A) allele have a 3.4-fold increased risk—adjusted odds ratio (OR) 3.42; $p = 0.013$—of lung cancer compared with wild-type individuals. Further, stratified analysis indicated that older individuals (>50 years) carrying the 723Gln (A) allele have the highest susceptibility to lung cancer (adjusted OR 4.10; $p = 0.020$). No substantial association was found between the Arg723Gln (2168G>A) polymorphism and chemotherapy response in Chinese lung cancer patients. The Arg723Gln (2168G>A) polymorphism of ABCC1 appears to be a potential susceptibility marker for lung cancer in the Chinese population, especially in older people [9].

44.3.2 ABCB1 and ABCC2

Some studies have found that ATP-binding cassette (ABC) efflux transporters play an important role in antiepileptic drug resistance, especially ABCB1 and ABCC2. We evaluated the relationship between the genetic polymorphisms of ABCC2 and ABCB1 and the therapeutic efficacy of antiepileptic drugs (AEDs) in Chinese epileptic patients. ABCB1 rs1045642 (3435C>T) and ABCC2 rs717620 (-24C>T), rs3740066 (3972C>T), and rs2273697 (1249G>A) polymorphism loci in 537 Chinese epilepsy patients (217 drug-resistant patients and 320 drug responders) were genotyped. ABCC2 rs717620-24TT genotype was found to be significantly associated with drug-resistant epilepsy (OR = 4.06, $p = 0.001$). The OR values of ABCC2 rs717620-24 CT+TT genotypes and ABCC2 rs3740066 (3972C>T) CT+TT genotypes were markedly higher in drug-resistant patients (OR = 1.57, $p = 0.018$; OR = 1.49, $p = 0.038$,

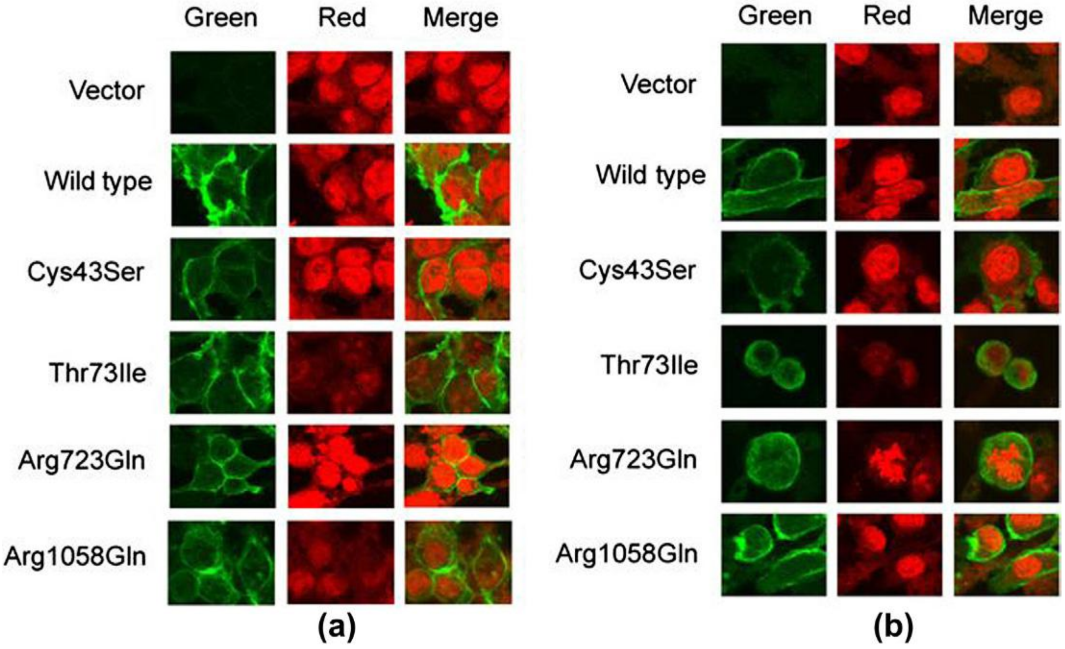


FIGURE 44.3 Confocal microscopy of wild-type and mutant ABCC1 expression in HEK293 and CHO-K1 cells. The subcellular localization of ABCC1 in HEK293(A) and CHO-K1(B) cells was detected by immunofluorescence. ABCC1's location is shown in green. Nuclei were stained with propidium iodide and are shown in red.

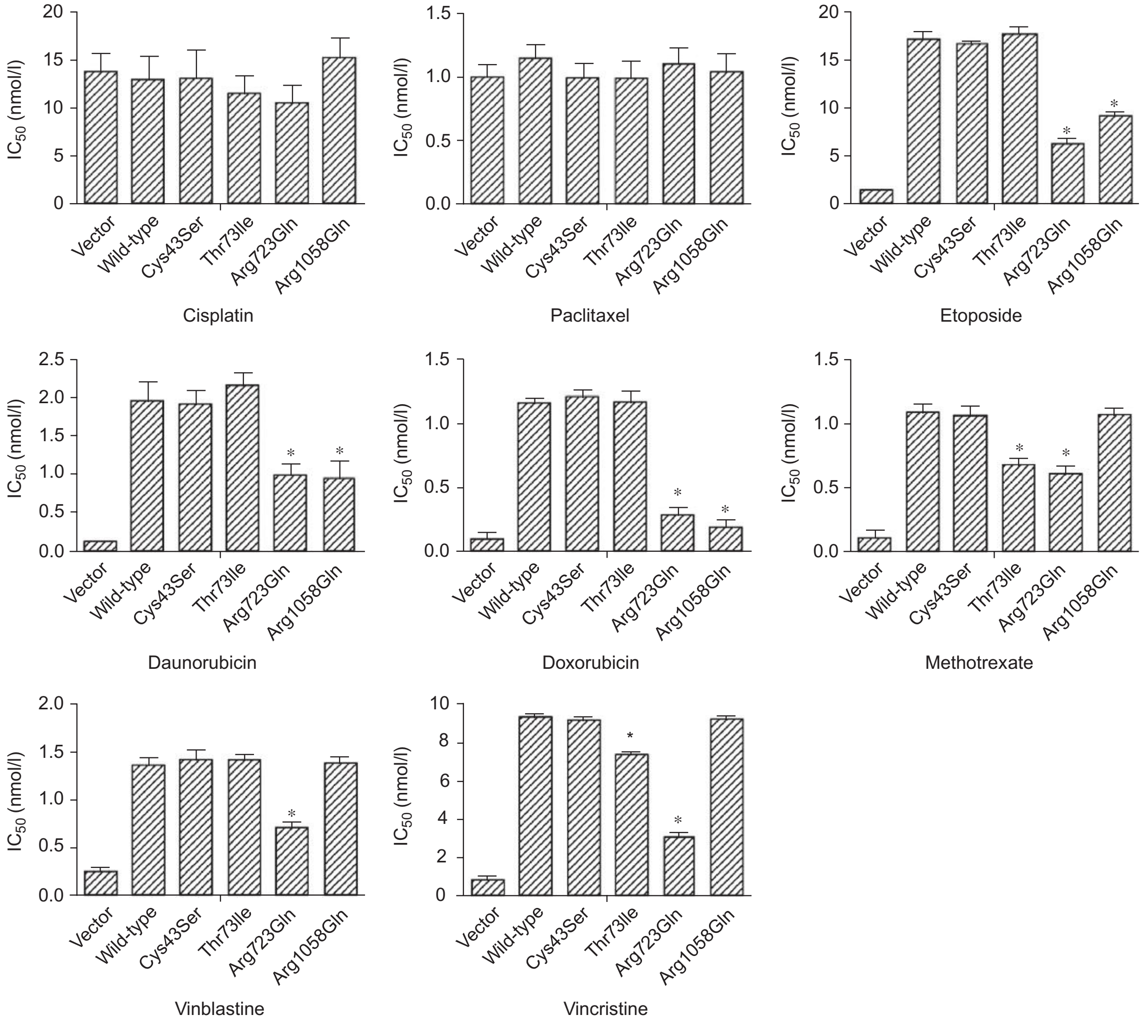


FIGURE 44.4 Effect of mutations on multidrug resistance-associated protein 1 (MRP1/ABCC1)-mediated multidrug resistance in HEK293 cells. HEK293 cells stably transfected with vector, wild-type, or Cys43Ser, Thr73Ile, Arg723Gln, and Arg1058Gln mutant MRP1/ABCC1 were exposed to cisplatin, paclitaxel, etoposide, daunorubicin, doxorubicin, methotrexate, vinblastine, and vincristine at various concentrations for 72 hours at 371 °C. Following was a methyl thiazolyl tetrazolium assay and determination of half maximal inhibitory concentration (IC50). The results shown are from six replicate determinations (*$p < 0.05$).

respectively) compared with responsive patients. ABCC2 rs2273697 (1249G>A) and ABCB1 rs1045642 (3435C>T) polymorphisms were not associated with drug-resistant epilepsy.

A linkage disequilibrium (LD) test showed that the ABCC2 rs717620 were in strong LD with rs2273697 ($D' = 0.694$) and rs3740066 ($D' = 0.699$). The frequencies of haplotypes TGT (ABCC2-24C>T/ABCC2 1249G>A/ABCC2 3972C>T) in resistant patients was significantly higher than those in responsive patients (21.0% vs. 14.2%, $p < 0.05$). These data suggest that ABCC2-24C>T, 3972C>T polymorphisms, and one ABCC2 haplotype, are associated with AED resistance and may affect the response to antiepileptic drugs [10].

44.3.3 ABCG2

Rosuvastatin, a novel, potent HMG-CoA reductase inhibitor, is excreted into bile mediated by breast cancer resistance protein (ABCG2). The association between the most frequent SNPs of the ABCG2 (421C>A) and the pharmacokinetics of rosuvastatin was studied in Chinese healthy volunteers. Prescreening of SLCO1B1 521TC and CYP2C9*1/*3 were performed before this pharmacokinetic study. Healthy volunteers who are SLCO1B1 521TT and CYP2C9*1/*1 wild-type homozygotes were selected to participate.

The pharmacokinetic parameters of rosuvastatin showed a significant difference between the two ABCG2 (421C>A) genotyped groups. The AUC (0–72) and AUC (0–) of rosuvastatin were much lower in the 421CC wild genotype group than in the mutation carriers (33.8 ± 11.4 vs. 59.6 ± 22.2 ng.h/ml, p = 0.018; 34.9 ± 11.9 vs. 62.2 ± 23.5 ng.h/ml, p = 0.018), respectively. The C(max) value was higher in the 421CA+421AA group than in the 421CC group (9.9 ± 5.4 vs. 5.1 ± 2.4 ng/ml, p = 0.048). The CL/F value was lower in the 421CA+421AA group than in the 421CC group (384.7 ± 161.2 vs. 674.0 ± 297.6 l/h, p = 0.043) (Figure 44.5).

The ABCG2 421C>A polymorphism may play an important role in the pharmacokinetics of rosuvastatin in healthy Chinese males after the exclusion of the impact of SLCO1B1 and CYP2C9 genetic polymorphisms. This was the first study on the association between ABCG2 polymorphism and rosuvastatin pharmacokinetics, which was further verified in a Finnish population and a Hong Kong Chinese population. Both PK and the low-density lipoprotein cholesterol response to rosuvastatin are associated with the ABCG2 421C>A polymorphism [11].

44.3.4 SLCO1B1

Organic anion transporting polypeptide 1B1 (OATP1B1) is a sodium-independent bile acid transporter, also known as liver-specific transporter 1 (LST-1) or OATP-C, encoded gene SLCO1B1. It is specifically expressed at the basolateral membrane of human hepatocytes and is responsible for the hepatocellular uptake of a variety of endogenous and foreign chemicals. A number of SNPs have been identified in the human SLCO1B1 gene. Our study clarified that common genetic polymorphisms occurred on OATP1B1 which had a potential impact on OATP1B1 function in Chinese populations.

For the first time, we found that the frequencies of the SLCO1B1 388G>A and 521T>C variant alleles in the Chinese population were 73.4% and 14.0%, respectively, which was close to that in the Japanese and Korean populations. The SLCO1B1*1a, SLCO1B1*1b, and SLCO1B1*15 are the three main haplotypes of SLCO1B1 in the Chinese population.

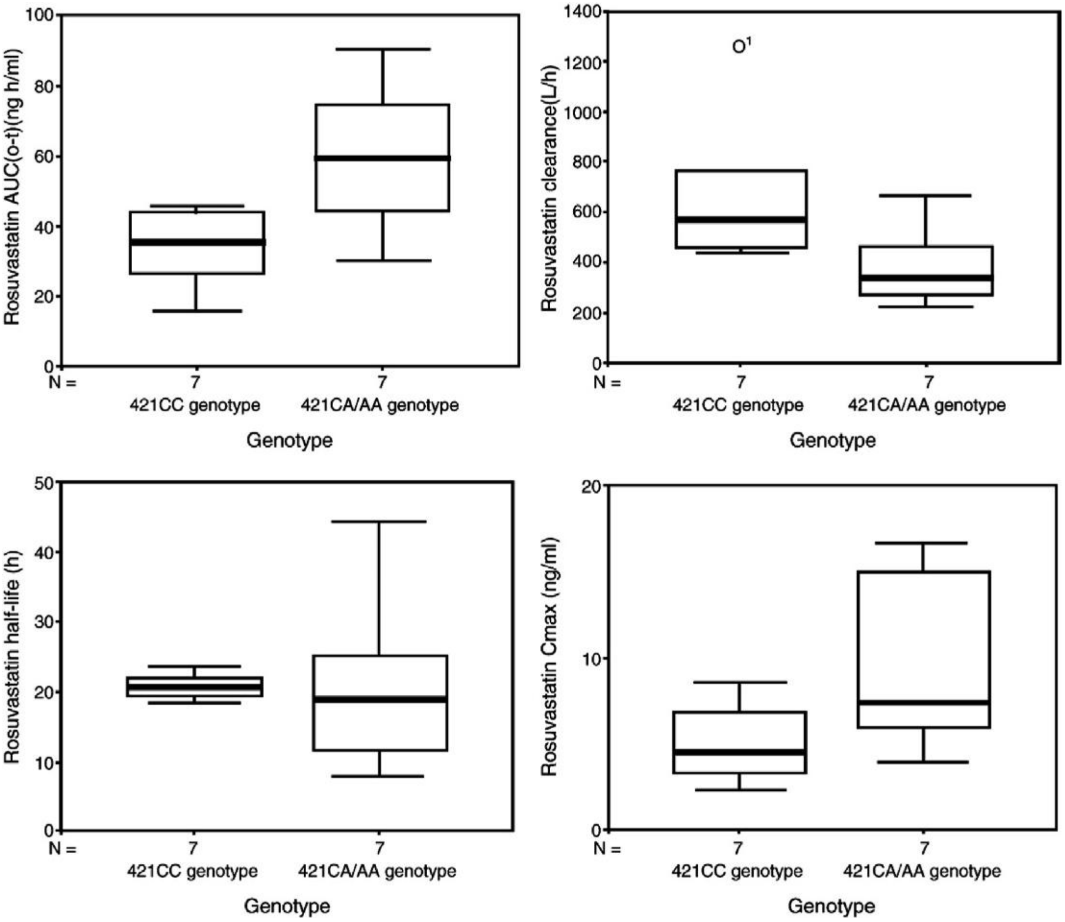


FIGURE 44.5 BCRP 421C>A genotype and the main pharmacokinetic parameters of a single dose of 20 mg rosuvastatin. Bars represent interquartile range; circles, outliers; lines across bars, mean. BCRP 421CC genotype group (n = 7), BCRP 421CA+421AA genotype group (n = 7).

Genetic polymorphism in SLCO1B1 is a major determinant of interindividual variability in serum bilirubin level. SLCO1B1 521T>C carriers had higher baseline serum UCB, T.BIL, and D.BIL levels compared with 521TT wild-type subjects. SLCO1B1 521CC subjects are more susceptible to hyperbilirubinemia. The serum bilirubin level could be increased by low-dose rifampicin administration, but the extent of elevation was not associated with the SLCO1B1 genotype.

We also found that OATP1B1-mediated hepatic uptake of nateglinide may be the prior step for its metabolism and elimination. SLCO1B1 521T>C SNP might play an important role in the pharmacokinetics of nateglinide. Total cholesterol level was significantly decreased after daily treatment with 20 mg pravastatin for 30 days in all patients. Moreover, subjects with the SLCO1B1 521C allele showed an attenuated total cholesterol-lowering effect compared with those homozygous for the 521T allele [12,13].

44.3.5 SLC30A8

Genome-wide association studies (GWASs) identified that SLC30A8 (zinc transporter) genetic polymorphism was a risk in type 2 diabetes mellitus (T2DM) in several populations. We investigated the association between SLC30A8 rs13266634 and rs16889462 polymorphisms and T2DM susceptibility and repaglinide therapeutic efficacy in Chinese T2DM patients. We conducted a case-control study of 443 T2DM patients and 229 healthy volunteers. The SLC30A8 rs13266634 risk C allele frequency was found to be higher in T2DM patients than in healthy controls ($P < 0.05$).

There was a better repaglinide response on FINS ($p < 0.05$) and PINS ($p < 0.01$) in patients with rs13266634 CT+TT genotypes compared with SLC30A8 CC genotype carriers. Patients with SLC30A8 rs16889462 GA genotype showed an enhanced repaglinide efficacy on FPG ($p < 0.01$), PPG ($p < 0.01$), and HbAlc ($p < 0.05$) compared with GG genotype individuals [14]. SLC30A8 rs13266634 and rs16889462 polymorphisms were associated with repaglinide therapeutic efficacy in Chinese T2DM patients.

44.4 PHARMACOGENOMICS OF DRUG RECEPTORS IN CHINESE POPULATIONS

44.4.1 β1-Adrenergic Receptor

The human β1-adrenergic receptor, an important therapeutic target in cardiovascular diseases, has two common functional polymorphisms (Ser49Gly and Gly389Arg). The distribution characteristic was clarified in Chinese Han healthy and those with essential hypertension. The allele frequencies of β1-adrenergic receptor 49Gly and 389Arg were 15.8% and 71.7% in healthy subjects, respectively. The effect of Gly389 polymorphism of the β1-adrenergic receptor on blood pressure and heart rate response to metoprolol was analyzed in a Chinese healthy population.

Exercise-induced heart rate increases were compared to determine the functional significance in vivo in 8 healthy Chinese men homozygous for Gly389 and 8 homozygous for Arg389. All of the subjects were given 25, 50, or 75 mg of metoprolol every 8 hours; the dosages were given in a random order, and each dosage was given for one day. The degree of β-blockade was measured as the reduction in resting and exercise heart rates and blood pressure. Plasma metoprolol concentrations were measured by the use of HPLC-fluorescence detection. Exercise led to a workload-dependent increase in heart rate. Oral metoprolol caused significant dose-dependent decreases in both resting and exercise heart rates in both groups. The reductions in the resting heart rate at 3 dosage levels of metoprolol were 6.3% ± 0.8% versus 4.1% ± 0.7%, 10.1% ± 1.0% versus 6.2% ± 1.1%, and 14.4% ± 1.4% versus 10.9% ± 1.3% in homozygous Arg389 subjects and Gly389 subjects, respectively ($p = 0.008$).

We also found differences with respect to the exercise heart rate (8.9% ± 0.5%, 14.0% ± 0.9%, and 20.1% ± 1.5% in Arg389 subjects and 6.6% ± 0.7%, 11.7% ± 1.0%, and 16.4% ± 1.3% in Gly389 subjects; $p = 0.017$) and systolic pressure (5.9% ± 0.7%, 9.2% ± 1.0%, and 11.6% ± 1.2% in Arg389 subjects and 4.6% ± 0.5%, 6.0% ± 0.8%, and 9.9% ± 0.9% in Gly389 subjects; $p = 0.011$). However, the difference in the fall in diastolic pressure was not statistically significant ($p = 0.442$). The Arg389 variant of the β1-adrenergic receptor was associated with a greater response to metoprolol than to Gly389 in young male Chinese subjects. These data suggested that the β1-adrenergic receptor Gly389Arg polymorphism is of major functional importance in Chinese healthy people.

In essential hypertension, the allele frequencies for 49Gly and 389Arg were 17.0% and 79.5%, respectively. There were significant differences in both genotype and allele frequency of Gly389Arg polymorphism between healthy and essential hypertension, which suggested that Gly389Arg may be involved in the pathophysiology of essential hypertension. The Arg389 variant of β1-adrenergic receptor was associated with a greater response to metoprolol than to Gly389 in young male Chinese subjects, which suggested that β1-adrenergic receptor Gly389Arg polymorphism is of major functional importance in vivo.

Therefore, we investigated the role of β1-adrenergic receptor polymorphisms on the blood pressure response to metoprolol monotherapy in the Chinese population with hypertension. Patients with certain β1-adrenergic receptor diplotypes—18 for 49Ser389Arg/49Ser389Arg, 15 for 49Ser389Arg/49Gly389Arg, 19 for 49Ser389Gly/49Gly389Arg, and 9 for 49Ser389Gly/49Ser389Gly—were selected for measurement of the antihypertensive effect of metoprolol. Patients were given 25 mg metoprolol every 12 hours for 4 weeks. The descent of systolic blood pressure after metoprolol administration was significantly different among genotype groups (10.4 ± 4.0%, 2.8 ± 4.7%, and 1.1 ± 1.5% for Arg389Arg, Gly389Arg, and Gly389Gly patients, respectively; p <0.001).

We also found a similar difference in changes of diastolic blood pressure (6.1 ± 4.3%, 2.2 ± 4.2%, and 0.9 ± 4.0%, respectively; p <0.001) and mean arterial pressure (8.1 ± 3.5%, 2.5 ± 3.0%, and 1.0 ± 2.5%, respectively; p <0.001) for Arg389Arg, Gly389Arg, and Gly389Gly patients. Ser49Gly variance exhibited a smaller contribution to the antihypertensive effect of metoprolol. Systolic blood pressure decreased significantly in Ser49 homozygous patients compared with Ser49Gly patients (8.4 ± 3.2% vs. 5.3 ± 5.2%; p = 0.047). There was a highly significant relationship between diplotype and blood pressure during treatment. Systolic blood pressure significantly decreased in 49Ser389Arg/49Ser389Arg (12.0 ± 3.8%, p <0.001) and 49Ser389Arg/49Gly389Arg (8.4 ± 5.5%, p <0.001) patients, with the decrease in the former being more pronounced (p = 0.023).

We also found a significant decrease in diastolic blood pressure (6.5 ± 4.7% vs. 5.7 ± 3.2%, respectively; both p <0.001) and mean arterial pressure (8.8 ± 3.2% vs. 6.9 ± 3.7%, respectively; both p <0.001) in 49Ser389Arg/49Ser389Arg and 49Ser389Arg/49Gly389Arg patients (Figure 44.6). β1-adrenergic receptor polymorphism was associated with different blood pressure responses to metoprolol therapy in patients with essential hypertension. 49Ser389Arg/49Ser389Arg and 49Ser389Arg/49Gly389Arg patients were good responders to metoprolol therapy; 49Ser389Arg/49Ser389Arg patients had a larger systolic blood pressure reduction than did 49Ser389Arg/49Gly389Arg patients. 49Ser389Gly/49Gly389Arg and 49Ser389Gly/49Ser389Gly patients were nonresponders to metoprolol antihypertensive therapy [15].

44.4.2 β3-Adrenergic rReceptor

β3-adrenergic receptor (ADRB3) belongs to the G protein-coupled receptor superfamily. It plays an important role in energy metabolism, which includes enhancing lipolysis, mobilization of free fatty acids into the circulation in white adipose tissue and thermogenesis in brown adipose tissue. ADRB3 is a potential target for antiobesity and antidiabetes therapy. We investigated the association between two novel ADRB3 gene polymorphisms (Ser165Pro and Ser257Pro) and type 2 diabetes (T2DM) in a Chinese population.

A total of 650 patients with T2DM and 1337 healthy volunteers were enrolled. Two candidate polymorphisms were recreated by site-directed mutagenesis and tested for their effect on ADRB3 expression and function in stable transfected human embryonic kidney 293 cells and Chinese hamster ovary-K1 cells. We found that both polymorphisms were significantly associated with T2DM (OR = 2.060 for Ser165Pro and 7.588 for Ser257Pro). T2DM patients with the Ser165Pro C allele had a higher hemoglobin A1c, fasting

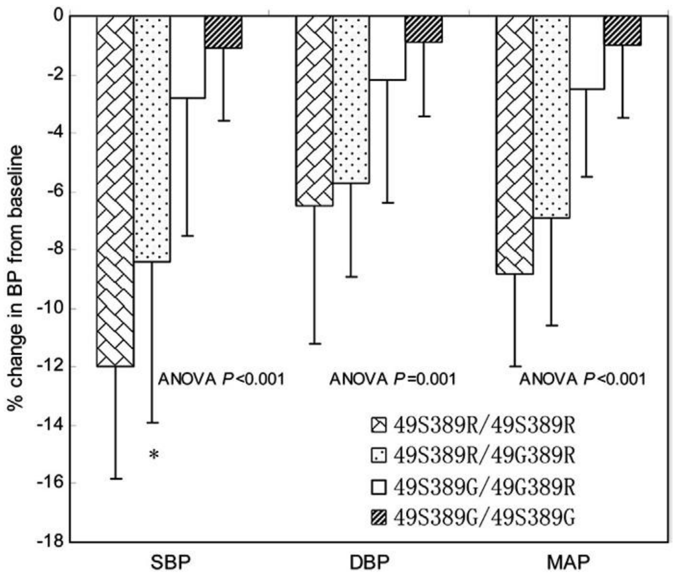


FIGURE 44.6 Blood pressure response to metoprolol monotherapy in patients with hypertension stratified according to β1-adrenergic receptor diplotypes. 49S389R/49S389R, n = 18; 49S389R/49G389R, n = 15; 49S389G/49G389R, n = 19; 49S389G/49S389G, n = 9. Data are presented as mean percentage decrease with SD. SBP—systolic blood pressure; DBP—diastolic blood pressure; MAP—mean arterial pressure. Asterisk denotes p = 0.023, compared with 49S389R/49S389R patients.

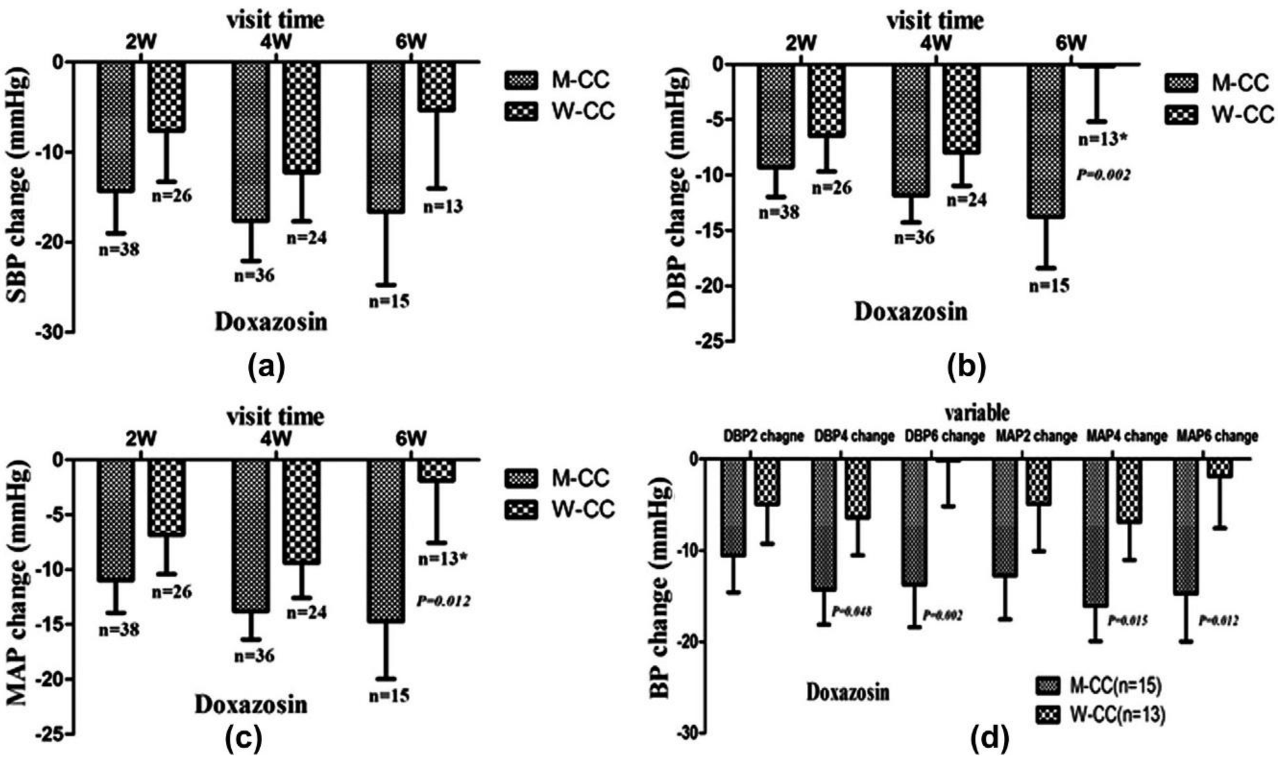


FIGURE 44.7 **KCNH2 (1956, C>T) genetic polymorphism in response to doxazosin is gender specific.** (a–c) depicts changes in systolic blood pressure, diastolic blood pressure, and mean arterial pressure in male CC genotype carriers versus female CC genotype carriers with essential hypertension after 2, 4, and 6 weeks of follow-up of those medicated with doxazosin. (d) No follow-up records were missed at the end of 2,4 and 6 weeks; the KCNH2(C195T) polymorphism related to doxazosin effect is gender-specific, *p <0.05 as compared with corresponding male CC carriers.

plasma glucose, and postprandial plasma glucose values than those in TT genotypes. We also found that patients with T2DM with the Ser257Pro C allele had lower fasting serum insulin, postprandial serum insulin, and homeostasis model assessment for insulin resistance levels than TT genotype carriers [16]. Further in vitro study indicated that cell lines stably expressing Ser165Pro and Ser257Pro mutants of the ADRB3 gene showed impaired cAMP accumulation activity.

44.4.3 PPAR-γ2

Pioglitazone is a prescription drug of the class thiazolidinedione (TZD) with hypoglycemic (antihyperglycemic, antidiabetic) action to treat diabetes. The influence of peroxisome proliferator-activated receptor γ2 (PPAR-γ2) gene polymorphism rs1801282 on the occurrence of type 2 diabetes and pioglitazone efficacy were studied in a Chinese Han population of 197 type 2 diabetes patients and 212 healthy controls. Among them, 67 T2DM patients were administered pioglitazone (30 mg/d, po) for three months. The PPAR-γ2 gene rs1801282 polymorphism was significantly associated with type 2 diabetes susceptibility (OR = 0.515). After pioglitazone treatment for three months, the T2DM patients with PPAR-γ2 rs1801282 CG genotypes showed significantly higher differential values of postprandial plasma glucose and serum triglyceride compared to those with rs1801282 CC genotype. Therefore, the diabetes risk allele in PPAR-γ2 (rs1801282) is also associated with pioglitazone therapeutic efficacy [17].

44.4.4 GABA Receptor

Carbamazepine (CBZ) is an anticonvulsant and mood-stabilizing drug used primarily in the treatment of epilepsy and bipolar disorder. The association between polymorphisms in gamma-aminobutyric acid (GABA) receptor genes and the tolerability and efficacy of carbamazepine monotherapy was investigated in 448 patients with partial-onset seizures. The effectiveness of CBZ treatment was evaluated with regard to efficacy by the decrease in seizures and tolerability by retention rates. Significant differences existed in the retention rates of CBZ over the 24 months of clinical follow-up with respect to the rs2290732 A/G polymorphism of the GABRA1 gene by Kaplan–Meier analysis (p = 0.001). For the CBZ treatment period, the overall retention rate was 50.7%. For the AA genotype, it was 62.9%; for the AG genotype 52.5%; and for the AA genotype only 28.2%. The rs2290732 A/G polymorphism of the GABRA1 gene might affect CBZ tolerability [18].

44.5 PHARMACOGENOMICS OF ION CHANNELS IN CHINESE POPULATIONS

44.5.1 KCNH2

KCNH2 (hERG) potassium channels have an integral role in regulating the excitability of smooth muscle cells. Some pathways driven by angiotensin II, nitric oxide, and adrenergic receptor blockers are involved in modulating the properties of KCNH2 potassium channels. These pathways are closely related to blood pressure regulation. Therefore, we hypothesized that KCNH2 genetic polymorphisms may affect blood pressure response to antihypertensive drug therapies. A total of 370 subjects with essential hypertension (EH) were studied to evaluate the interactions between KCNH2 genetic polymorphisms and individual blood pressure response to antihypertensive drugs. There were statistically significant interactions between KCNH2 (1956, C>T) polymorphism and DBP change ($p = 0.010$), and MAP change ($p = 0.014$) on azelnidipine or nitrendipine therapy patients at the end of six weeks. We found that the KCNH2 (1956, C>T) polymorphism was associated with the hypotensive effects of α,β-ADR blockers of DBP change at the end of four and six weeks of treatment in an age- and gender-dependent manner ($p = 0.007$ and 0.019, respectively).

Similar results were also observed for changes in MAP at the end of four and six weeks ($p = 0.035$ and 0.078, respectively). It was the first time that the KCNH2 (1956, C>T) polymorphism was found to be associated with the efficacy of antihypertensive CCBs and ADR blockers, and it may serve as a novel biomarker for individualized therapy for these antihypertensive drugs (Figure 44.7). We subsequently summarized the current pharmacogenomic studies on hERG potassium channels to provide a reference for developing individualized treatments and evaluating their safety. The genetic polymorphisms in hERG potassium channels are associated with many complex diseases and sensitivity to channel-related drugs. Genotypes may underlie different sensitivities to the same drug, and different drugs selectively repair the functional deficits caused by individual mutations (Figure 44.8).

In fact, not all drugs that block the hERG function have adverse affects, as previously thought. This suggests that the

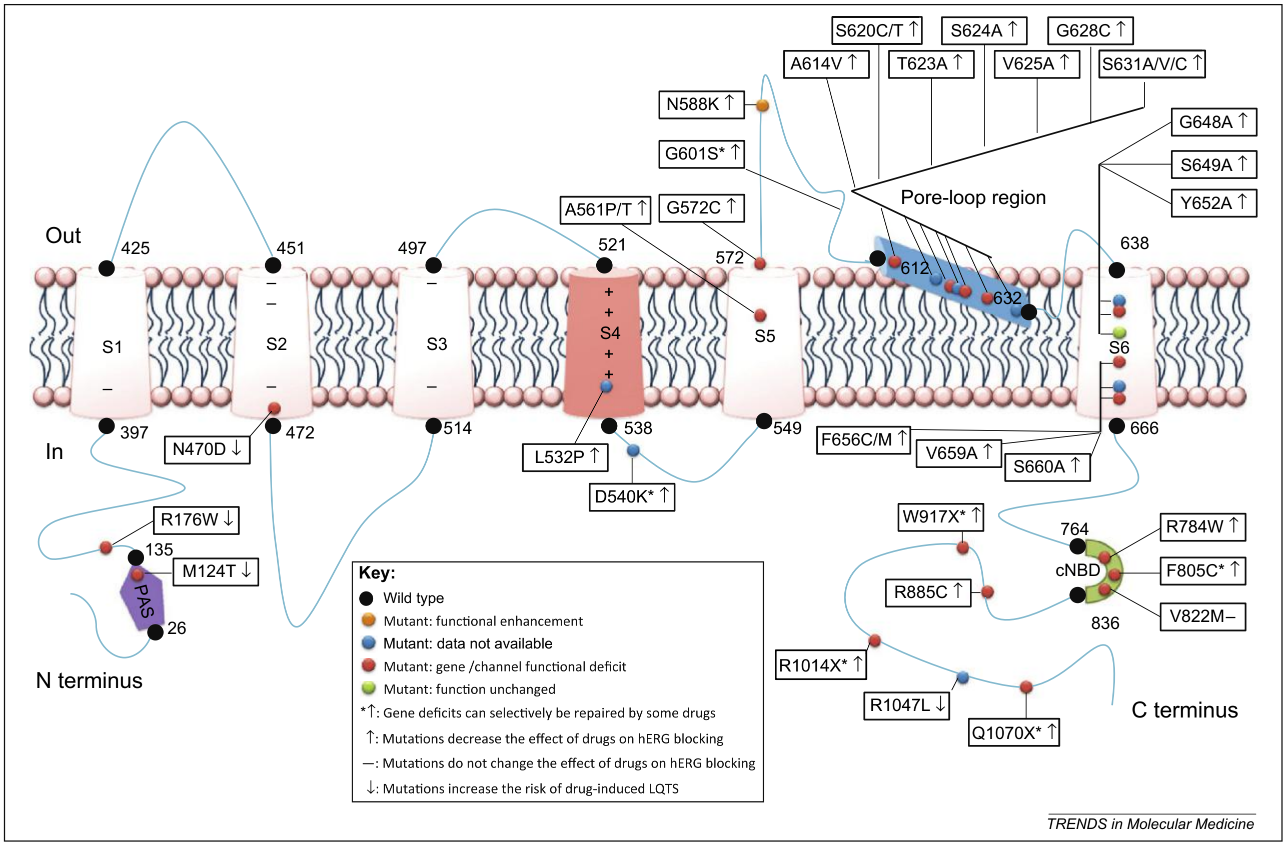

S620C/T ↑
S624A ↑
G628C ↑
A614V ↑
T623A ↑
V625A ↑
S631A/V/C ↑
N588K ↑
G648A ↑
G601S* ↑
S649A ↑
Pore-loop region
A561P/T ↑
G572C ↑
Y652A ↑
Out
425
451
497
521
572
612
632
638
S1
S2
S3
S4
S5
S6
397
472
514
538
549
666
In
N470D ↓
L532P ↑
F656C/M ↑
V659A ↑
S660A ↑
D540K* ↑
R176W ↓
W917X* ↑
764
R784W ↑
135
M124T ↓
cNBD
F805C* ↑
PAS
R885C ↑
26
836
V822M–
N terminus
R1014X* ↑
R1047L ↓
Q1070X* ↑
C terminus
Key:
● Wild type
● Mutant: functional enhancement
● Mutant: data not available
● Mutant: gene /channel functional deficit
● Mutant: function unchanged
*↑: Gene deficits can selectively be repaired by some drugs
↑: Mutations decrease the effect of drugs on hERG blocking
—: Mutations do not change the effect of drugs on hERG blocking
↓: Mutations increase the risk of drug-induced LQTS
TRENDS in Molecular Medicine

FIGURE 44.8 **Human hERG structure and polymorphisms.** Shown are a single hERG subunit containing 6-α helical transmembrane domains (S1–S6), a pore-forming loop (between S5 and S6), a C-terminal cyclic nucleotide-binding domain (cNBD), and N-terminal Per-Arnt-Sim (PAS) domains. Thirty KCNH2 SNPs are associated with hERG potassium channel sensitivity to about 70 drugs; polymorphisms leading to changes in channel properties are highlighted.

severe adverse reactions observed clinically may only occur in subjects with a certain genotype; others may be safe. Similarly, a drug that is ineffective in one population may be both safe and effective in another. Therefore, detecting KCNH2 polymorphisms is of great significance in guiding the prevention and treatment of related diseases, reevaluating drug safety, and individualizing treatment [19,20].

44.5.2 KCNQ1

KCNQ1 is a voltage-gated potassium channel present mainly in the cell membranes of cardiac tissue and in inner-ear neurons, among other tissues. We evaluated the effects of KCNQ1 rs2237892 and rs2237895 polymorphisms on repaglinide efficacy in Chinese patients with type 2 diabetes mellitus (T2DM). In all, 367 T2DM patients and 214 controls were genotyped. Forty of the T2DM patients were randomly selected to undergo 8 weeks of repaglinide treatment. The frequency of the rs2237892 allele was lower in the T2DM patients than in the control group ($p < 0.05$). The frequency of the rs2237895 C allele was higher in T2DM patients than in healthy control subjects ($p < 0.05$). Diabetic patients with the rs2237892 risk C allele had lower fasting insulin levels ($p < 0.01$) and homeostasis model assessment of insulin resistance (HOMA-IR; $p < 0.01$) values than carriers of the T allele.

Diabetic patients with the rs2237895 risk C allele had higher fasting plasma glucose ($p < 0.01$), postprandial plasma glucose (PPG) levels ($p < 0.01$) and HOMA-IR values ($p < 0.01$) than those with the A allele. Following repaglinide treatment, those T2DM patients with the rs2237892 T allele and rs2237895 C allele were more likely to have a positive response to repaglinide in terms of PPG levels ($p < 0.05$) than T2DM patients with the rs2237892 CC and rs2237895 AA genotypes [21]. Therefore, KCNQ1 rs2237892 and rs2237895 polymorphisms were associated with the therapeutic efficacy of repaglinide in Chinese T2DM patients.

44.5.3 KCNJ11

Kir6.2 is a major subunit of the ATP-sensitive K+ channel, an inward-rectifier potassium ion channel. The gene encoding the channel is called KCNJ11, and mutations in this gene are associated with congenital hyperinsulinism. We have shown that the polymorphism Lys23Glu is associated with a heightened risk of developing T2DM. We also explored the effects of the polymorphism on the efficacy of repaglinide therapy in Chinese patients with T2DM.

A total of 259 patients with T2DM and 188 healthy controls were genotyped. Forty patients with various genotypes were randomly selected to undergo an 8-week repaglinide treatment regimen. Patients with the G allele of the KCNJ11 Lys23Glu polymorphism showed higher levels of fasting plasma glucose (FPG) and postprandial plasma glucose (PPG) ($p < 0.05$). After repaglinide treatment, patients with the GA or AA genotype showed higher levels of FPG, PPG, and glycated hemoglobin (HbA(1c)) compared to patients with the GG genotype ($p < 0.05$) [22]. Thus, the KCNJ11 polymorphism was associated with repaglinide efficacy in this Chinese population.

44.5.4 SCN1A

The SCN1A gene belongs to a family of genes that provide instructions for making sodium channels, which transport positively charged sodium atoms (sodium ions) into cells and play a key role in a cell's ability to generate and transmit electrical signals. We investigated the tolerability and efficacy of carbamazepine treatment in patients with partial-onset seizures and the association with polymorphisms in the sodium channel α-subunit type 1 (SCN1A) among 448 Chinese Han patients. Monotherapy with CBZ was administered to the patients. The effectiveness of CBZ treatment was evaluated with regard to efficacy by the decrease in seizures and tolerability by retention rates. SCN1A rs3812718 A/G with CBZ tolerability ($p = 0.038$) throughout a 24-month clinical follow-up. The maintenance dose and serum level of CBZ in AA genotype carriers of rs3812718 A/G was significantly higher than those of GG genotype carriers between 3 and 12 months. The proportion of AA genotype carriers of rs2298771 A/G that were seizure free was significantly higher than that for AG+GG genotype carriers from 3 months to 15 months of follow-up ($p < 0.05$). Thus, SCN1A rs3812718 A/G polymorphism affected the tolerability of CBZ and rs2298771 A/G was associated with efficacy of CBZ treatment [18].

We subsequently investigated whether SNPs in the SCN1A gene could affect the retention rate of carbamazepine used to treat seizures in Chinese Han patients with epilepsy. In total, 448 patients were genotyped for SNPs selected in the SCN1A gene. The tag SNPs were selected using Haploview version 4.2 software (www.broad.mit.edu/haploview/haploview). Monotherapy with CBZ was administered to patients with new-onset focal seizures. In the present study, the retention rate of CBZ was defined as the percentage of patients with epilepsy who had continued with the treatment during the preceding three months.

Potential confounding variables were analyzed to evaluate the risk factors for nonretention. When adjusted for potential confounding factors (e.g., age, sex, body mass index, seizure type), the presence of the A allele of SNP rs3812718 predicted nonretention ($p = 0.015$; $OR = 3.122$). Beyond 12 months of treatment in those who continued with CBZ (retention cohort), analysis of covariance indicated that the maintenance dose ($p = 0.025$) and serum CBZ levels ($p = 0.021$) were significantly higher in carriers of the rs3812718 AA genotype than in those with the

GG genotype [23]. The results of the present study indicate that the rs3812718 SNP in the SCN1A gene is significantly associated with the retention rate of CBZ monotherapy in Chinese Han patients with focal epilepsy.

44.6 ENVIRONMENTAL FACTORS IN DRUG-METABOLIZING ENZYME AND TRANSPORTER ACTIVITY

44.6.1 CYP2C19-Mediated Herb–Drug Interaction

St. John's wort, an extract of the medicinal plant *Hypericum perforatum*, is widely used as an herbal antidepressant. More recently, its ability to induce both CYP3A4- and CYP2C19-mediated reactions has been well established. Twelve Chinese healthy adult men (6 CYP2C19*1/CYP2C19*1, 4 CYP2C19*2/*2, and 2 CYP2C19*2/CYP2C19*3) were enrolled in a two-phase randomized crossover design. After a 14-day treatment with St. John's wort, substantial decreases in plasma concentrations of omeprazole were observed. The peak plasma concentration (C_{max}) significantly decreased by 37.5% ± 13.3% (p <0.001) in CYP2C19*2/*2 or *3 and by 49.6% ± 20.7% (p = 0.017) in CYP2C19*1/*1; the area under the concentration–time curve (AUC) extrapolated to infinity—AUC (0–∞)—decreased by 37.9% ± 21.3% (p = 0.014) and 43.9% ± 23.7% (p = 0.011) in CYP2C19 mutant and wild genotypes, respectively.

Moreover, the C_{max} and AUC (0–∞) of omeprazole sulfone increased by 160.3% ± 45.5% (p <0.001) and by 136.6% ± 84.6% (p = 0.014), 155.5% ± 58.8% (p <0.001), and 158.7% ± 101.4% (p = 0.017) in mutant and wild genotypes, respectively. St. John's wort increased the C_{max} of 5-hydroxyomeprazole by 38.1% ± 30.5% (p = 0.028) and the AUC (0–∞) by 37.2% ± 26% (p >0.005) in CYP2C19 wild-type subjects, whereas it did not produce any significant alterations to the corresponding pharmacokinetic parameters in subjects with variant genotypes. Clinically relevant interactions with other drugs may occur and must be taken into account when St. John's wort is being taken [24].

44.6.2 OATP1B1-Mediated Herb–Drug Interaction

Baicalin is a flavone glucuronide purified from the medicinal plant *Radix scutellariae* through uridine diphosphate glucuronidation. Nowadays, baicalin has begun to be used in bilirubin-lowering therapy, both prescribed and over the counter, in China. We explored the potential herb–drug interaction between baicalin and rosuvastatin, a typical substrate for organic anion-transporting polypeptide 1B1 (OATP1B1) related to different OATP1B1 haplotype groups.

Eighteen unrelated healthy volunteers who were CYP2C9*1/*1 with different OATP1B1 haplotypes (six OATP1B1*1b/*1b, six OATP1B1*1b/*15, and six OATP1B1*15/*15) were selected to participate in this study. Rosuvastatin (20 mg orally) pharmacokinetics after coadministration of placebo and 50-mg baicalin tablets (three times/day orally for 14 days) were measured for up to 72 h in a two-phase randomized crossover study. After baicalin treatment, the area under the plasma concentration–time curve (AUC) (0–72) and AUC (0–∞) of rosuvastatin decreased by 47.0 ± 11.0% (p = 0.001) and 41.9 ± 7.19% (p = 0.001) in OATP1B1*1b/*1b, 21.0 ± 20.6% (p = 0.035) and 23.9 ± 8.66% (p = 0.004) in OATP1B1*1b/*15, and 9.20 ± 11.6% (p = 0.077) and 1.76 ± 4.89% (p = 0.36) in OATP1B1*15/*15, respectively (Figure 44.9). Moreover, decreases of both AUC (0–72) and AUC (0–∞) of rosuvastatin among different haplotype groups were significantly different (p = 0.002 and <0.001) [25]. Thus, baicalin reduces plasma concentrations of rosuvastatin in an OATP1B1 haplotype-dependent manner.

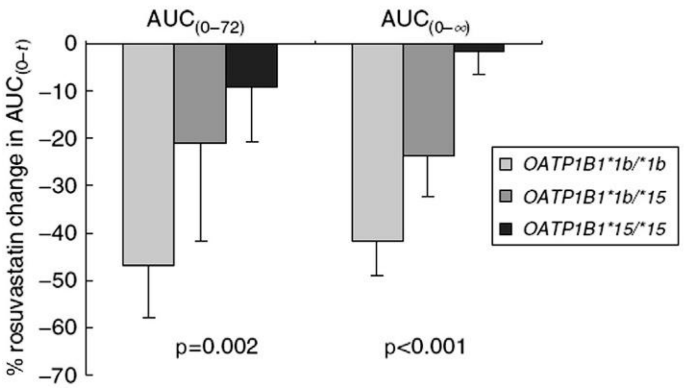


FIGURE 44.9 **Herb–drug interaction between rosuvastatin and baicalin is OATP1B1 genotype-related.** Comparison of mean change in the AUC (0–72) and AUC (0–∞) of rosuvastatin among *OATP1B1*1b/*1b*, *OATP1B1*1b/*15*, and *OATP1B1*15/*15* subjects after 14 days of treatment with baicalin compared with placebo. Data are expressed as mean ± SE.

44.7 TRANSLATIONAL APPROACHES TO PHARMACOGENOMICS IN CHINA

Translational medicine integrates research from the basic sciences, social sciences, and political sciences with the aim of optimizing patient care and preventive measures that may extend beyond healthcare services. The term was first used by the National Institutes of Health (NIH) in 2003 and was quickly picked up by the Chinese Natural Science Society. In 2006, AstraZeneca invested $100 million to fund an AstraZeneca Innovation Center and start translational medicine research in China; the aim is to develop Chinese-specific medicines.

In 2011, the first national translational medicine research center of China was funded in Shanghai; however, research on translational medicine had begun years before that. In 1993, Professor Hong-Hao Zhou funded China's first pharmacogenetics research institute. Contemporary common sense acknowledges the implementation of pharmacogenetics and genomics as the most successful and promising pathway of translational medicine. The work of Zhou is regarded as the actual start of translational medicine research in China. To date, more than nine translational research centers have been funded, including those at Peking University, Shanghai Institutes for Biological Sciences, Fudan University, Shanghai Jiao Tong University, Zhe Jiang University, Central South University, Wen Zhou University, Xia Men University, Guang Zhou Cancer hospital, and so on.

In the United States, the NIH and FDA play vital roles in sponsoring translational medicine research and keep expediting the translation of research from bench side to bedside. In China, institutions receive research funds from several different departments including the Ministry of Science and Technology, Ministry of Education, Chinese Academy of Sciences, and others. Key projects include the National High-tech R&D Program of China (863 Program), National Program on Key Basic Research Project (973 Program), National Basic Research Priorities Program of China, National Key Technologies R & D Program of China, National Natural Science Foundation of China (General Program; Key Program; Major Program), among others.

The Chinese government is spending millions of dollars each year to accelerate research and implementation in translational medicine. In addition, the whole field is overseen and regulated by the Ministry of Health (MOH). In the past few years, several companies have started businesses in pharmacogenomic molecular diagnosis and tried to provide personalized drug therapy via genotyping patients for certain genetic biomarkers. However, lack of legal permits has led to low-quality service from most companies. In 2013, MOH decided to issue the first regulations and standard operation procedures (SOP) for all laboratories involved in the clinical service of personalized medicine. Professor Hong-Hao Zhou's group has been selected to compile the regulations and SOPs, and is in charge of training and education for the nation. In addition, three model laboratories were selected: Xiang Ya Medical Lab Central South University, CapitalBio Corporation, and the first hospital of China Medical University to adopt the regulations and SOPs.

Pharmacogenomic translational research is still in an early phase in China, with many problems in application of the research findings. For instance, basic research needs to be strengthened and genetic markers that can definitely predict a drug's safety and efficacy need to be identified by comprehensive analysis based on pharmacokinetics and pharmacodynamics, efficacy, and toxicity as well. More attention should be given to translational research findings during the development of new drugs, and the clinical models of genetic tests for personalized medicine should be improved. In addition, education and training systems should be established for community health departments, medical institutions, medical students, doctors, and patients to promote public acceptance of personalized medicine.

44.8 CONCLUSION AND FUTURE PERSPECTIVES

The ideas of pharmacogenomics and personalized medicine have been accepted worldwide by medical institutions, drug administration agencies, and pharmaceutical companies; moreover, key technologies have gradually reached maturity with the support of major national programs. Thus, pharmacogenetics and pharmacogenomics will play more and more important roles in translational medicine and rational drug administration. It is believed that in the near future translational medicine and personalized medicine will become a major force in revolutionizing modern medicine in China and worldwide.

ACKNOWLEDGMENTS

This series of work was supported by the National Scientific Foundation of China (No. 30801421, 81273595) and the State High-Tech Development Plan of China (863 Project Nos. 2012AA02A518 and 2012AA02A517).

REFERENCES

[1] Zhou HH, Koshakji RP, Silberstein DJ, Wilkinson GR, Wood AJ. Altered sensitivity to and clearance of propranolol in men of Chinese descent as compared with American whites. N Engl J Med 1989;320(9):565–70.

[2] Xiao ZS, Goldstein JA, Xie HG, Blaisdell J, Wang W, Jiang CH, et al. Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare CYP2C19 mutant allele. J Pharmacol Exp Ther 1997;281(1):604–9.

[3] Shu Y, Wang LS, Xu ZH, He N, Xiao WM, Wang W, et al. 5-hydroxylation of omeprazole by human liver microsomal fractions

from Chinese populations related to CYP2C19 gene dose and individual ethnicity. J Pharmacol Exp Ther 2000;295(2):844–51.

[4] Qin XP, Xie HG, Wang W, He N, Huang SL, Xu ZH, et al. Effect of the gene dosage of CYP2C19 on diazepam metabolism in Chinese subjects. Clin Pharmacol Ther 1999;66(6):642–6.

[5] Hu YF, He J, Chen GL, Wang D, Liu ZQ, Zhang C, et al. CYP3A5*3 and CYP3A4*18 single nucleotide polymorphisms in a Chinese population. Clin Chim Acta 2005;353(1–2):187–92.

[6] Chen GL, Wang H, Wang W, Xu ZH, Zhou G, He F, et al. Histamine N-methyltransferase gene polymorphisms in Chinese and their relationship with enzyme activity in erythrocytes. Pharmacogenetics 2003;13(7):389–97.

[7] Chen GL, Zhu B, Nie WP, Xu ZH, Tan ZR, Zhou G, et al. Single nucleotide polymorphisms and haplotypes of histamine N-methyltransferase in patients with gastric ulcer. Inflamm Res 2004;53(9):484–8.

[8] Yin JY, Huang Q, Yang Y, Zhang JT, Zhong MZ, Zhou HH, et al. Characterization and analyses of multidrug resistance-associated protein 1 (MRP1/ABCC1) polymorphisms in Chinese population. Pharmacogenet Genomics 2009;19(3):206–16.

[9] Yin JY, Han LF, Huang Q, Xu XJ, Zhou HH, Liu ZQ. ABCC1 polymorphism Arg723Gln (2168G>A) is associated with lung cancer susceptibility in a Chinese population. Clin Exp Pharmacol Physiol 2011;38(9):632–7.

[10] Qu J, Zhou BT, Yin JY, Xu XJ, Zhao YC, Lei GH, et al. ABCC2 polymorphisms and haplotype are associated with drug resistance in Chinese epileptic patients. CNS Neurosci Ther 2012;18(8):647–51.

[11] Zhang W, Yu BN, He YJ, Fan L, Li Q, Liu ZQ, et al. Role of BCRP 421C>A polymorphism on rosuvastatin pharmacokinetics in healthy Chinese males. Clin Chim Acta 2006;373(1–2):99–103.

[12] Zhang W, Chen BL, Ozdemir V, He YJ, Zhou G, Peng DD, et al. SLCO1B1 521T>C functional genetic polymorphism and lipid-lowering efficacy of multiple-dose pravastatin in Chinese coronary heart disease patients. Br J Clin Pharmacol 2007;64(3):346–52.

[13] Zhang W, He YJ, Han CT, Liu ZQ, Li Q, Fan L, et al. Effect of SLCO1B1 genetic polymorphism on the pharmacokinetics of nateglinide. Br J Clin Pharmacol 2006;62(5):567–72.

[14] Huang Q, Yin JY, Dai XP, Wu J, Chen X, Deng CS, et al. Association analysis of SLC30A8 rs13266634 and rs16889462 polymorphisms with type 2 diabetes mellitus and repaglinide response in Chinese patients. Eur J Clin Pharmacol 2010;66(12):1207–15.

[15] Liu J, Liu ZQ, Tan ZR, Chen XP, Wang LS, Zhou G, et al. Gly389Arg polymorphism of β1-adrenergic receptor is associated with the cardiovascular response to metoprolol. Clin Pharmacol Ther. 2003;74(4):372–9.

[16] Huang Q, Yang TL, Tang BS, Chen X, Huang X, Luo XH, et al. Two novel functional single nucleotide polymorphisms of ADRB3 are associated with type 2 diabetes in the Chinese population. J Clin Endocrinol Metab 2013;98(7):E1272–7.

[17] Pei Q, Huang Q, Yang GP, Zhao YC, Yin JY, Song M, et al. PPAR-γ2 and PTPRD gene polymorphisms influence type 2 diabetes patients' response to pioglitazone in China. Acta Pharmacol Sin 2013;34(2):255–61.

[18] Zhou BT, Zhou QH, Yin JY, Li GL, Qu J, Xu XJ, et al. Effects of SCN1A and GABA receptor genetic polymorphisms on carbamazepine tolerability and efficacy in Chinese patients with partial seizures: 2-year longitudinal clinical follow-up. CNS Neurosci Ther 2012;18(7):566–72.

[19] He FZ, Luo J, Luo Z, Fan L, He Y, Zhu D, et al. The KCNH2 Genetic Polymorphism (1956, C>T) Is a Novel Biomarker That Is Associated with CCB and α, β-ADR Blocker Response in EH Patients in China. PLoS One 2013;8(4):e61317.

[20] He FZ, McLeod HL, Zhang W. Current pharmacogenomic studies on hERG potassium channels. Trends Mol Med 2013;19(4):227–38.

[21] Dai XP, Huang Q, Yin JY, Guo Y, Gong ZC, Lei MX, et al. KCNQ1 gene polymorphisms are associated with the therapeutic efficacy of repaglinide in Chinese type 2 diabetic patients. Clin Exp Pharmacol Physiol 2012;39(5):462–8.

[22] Yu M, Xu XJ, Yin JY, Wu J, Chen X, Gong ZC, et al. KCNJ11 Lys23Glu and TCF7L2 rs290487(C/T) polymorphisms affect therapeutic efficacy of repaglinide in Chinese patients with type 2 diabetes. Clin Pharmacol Ther 2010;87(3):330–5.

[23] Zhou BT, Zhou QH, Yin JY, Li GL, Xu XJ, Qu J, et al. Comprehensive analysis of the association of SCN1A gene polymorphisms with the retention rate of carbamazepine following monotherapy for new-onset focal seizures in the Chinese Han population. Clin Exp Pharmacol Physiol 2012;39(4):379–84.

[24] Wang LS, Zhou G, Zhu B, Wu J, Wang JG, Abd El-Aty AM, et al. St John's Wort induces both cytochrome P450 3A4-catalyzed sulfoxidation and 2C19-dependent hydroxylation of omeprazole. Clin Pharmacol Ther 2004;75(3):191–7.

[25] Fan L, Zhang W, Guo D, Tan ZR, Xu P, Li Q, et al. The effect of herbal medicine baicalin on pharmacokinetics of rosuvastatin, substrate of organic anion-transporting polypeptide 1B1. Clin Pharmacol Ther 2008;83(3):471–6.

This page is intentionally left blank

Chapter 45

Pharmacogenomics in Brazil

Guilherme Suarez-Kurtz[a] and **Mara Helena Hutz**[b]

[a]*Pharmacology Division, Instituto Nacional de Câncer, Brazil,* [b]*Department of Genetics, Universidade Federal do Rio Grande do Sul, Brazil*

Chapter Outline

45.1 Overview 1015
45.2 The Brazilian Pharmacogenetics Network–Refargen 1015
  45.2.1 Pharmacogenomic Diversity among Brazilians 1017
  45.2.2 PGx Clinical Trials in Brazilians 1021
45.3 PGX of Immunosuppressants in Renal Transplant Patients 1028
45.4 PGX of Antipsychotics in Schizophrenic Patients 1030
45.5 Conclusions 1032
Acknowledgments 1033
References 1033

45.1 OVERVIEW

The study of pharmacogenomics in Latin American countries must take into account ethnic, cultural, socioeconomical, scientific, and technological disparities, in addition to the genetic heterogeneity and population diversity across the region. The Portuguese language distinguishes Brazil from all other Latin American nations in which Spanish is the official idiom. Although Native American (Amerindian), European, and Sub-Saharan African ancestral roots are common to all of Latin America, the relative proportions of each root vary markedly across nations: for example, the average Amerindian ancestry is considerably higher in Mexico and Peru, compared to Brazil and Cuba, whereas the opposite prevails with respect to African ancestry. Population heterogeneity is also present across geographical regions, among and within "ethnic/race/Color" categories in the same country, Brazil providing particularly rich examples [1,2].

Based on gross domestic product (GDP), Brazil ranks as the 6th and Mexico as the 14th largest world economies, whereas their neighbors, Argentina and Guatemala, rank 26th and 76th, respectively (Table 45.1). However, a considerably different picture emerges from the GDP per capita, with Brazil, Mexico, Argentina, and Guatemala appearing in the 78th, 65th, 52th, and 122th positions, respectively. Scientific and technological production, as assessed by the number of indexed publications by country of origin, has Brazil, Mexico, Argentina, and Guatemala in the 13th, 29th, 36th, and 125th positions, respectively. Not surprisingly, the extent of pharmacogenetics/genomics (PGx) investigation varies considerably across Latin America; Brazil, the focus of the present chapter, has been the main source of PGx information in Latin American countries (Table 45.1).

45.2 THE BRAZILIAN PHARMACOGENETICS NETWORK–REFARGEN

Brazil is the fifth largest country in the world and occupies an area of 8.5 million km^2. Its present population, in excess of 200 million individuals, is highly heterogeneous and admixed, a fact that has far-reaching consequences for PGx [3–5]. Recognition of this fact prompted the creation of a nationwide network—the Rede Nacional de Farmacogenética or Refargen [6], which presently comprises 20 research groups distributed over four of the five geographical regions of Brazil. These groups are described in the Refargen Internet site [7], and their major research interests are listed in Table 45.2.

Among Refargen's initiatives in recent years, we will mention the coordination of the Second Latin American Pharmacogenomics and Personalized Medicine Congress attended by a broad audience, mostly from Brazil, but also from Argentina, Chile, Colombia, Costa Rica, Cuba, Ecuador, and Mexico. Leading Brazilian and international

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00045-1
© 2014 Elsevier Inc. All rights reserved.

TABLE 45.1 Population, Area, Economic Facts, and Scientific Output of Latin American Countries

Country	Population (million)[a]	Area (sq miles)[a]	GDP[b]		Per capita GDP[c]		Indexed Articles[d]			Indexed PGx Articles[e]
			US$B	*Rank*	*US$1000*	*Rank*	*Number*	*%*	*Rank*	*Number*
Argentina	41.8	1,068,296	485	26	18,112	52	7,272	0.61	36	5
Bolivia	10.1	424,162	24	98	5,099	123	203	0.02	108	0
Brazil	203.4	3,286,470	2,477	6	11,875	78	32,1	2.69	13	53
Chile	16.9	292,258	249	38	18,419	50	4,883	0.41	44	1
Colombia	44.7	439,733	333	31	10,792	82	2,47	0.21	53	5
Costa Rica	4.6	1,973	41	82	12,606	75	442	0.04	86	0
Cuba	11.1	42,803	69	63			814	0.07	70	3
Dominican Republic	9.9	18,815	55	71	9,646	91	37	<0.01	156	0
Ecuador	15.0	109,483	66	64	10,056	88	339	0.03	93	0
El Salvador	6.1	8,124	23	100	7,438	104	27	<0.01	163	0
Guatemala	13.8	42,042	47	76	5,209	122	103	0.01	125	0
Honduras	8.1	43,278	17	109	4,610	129	39	<0.01	155	0
Mexico	113.7	761,602	1,155	14	15,312	65	9,612	0.81	29	22
Nicaragua	5.7	49,998	7	138	4,458	130	58	<0.01	143	1
Panama	3.5	30,193	31	88	15,617	63	268	0.02	98	0
Paraguay	6.4	157,046	23	101	6,136	113	40	<0.01	154	0
Peru	29.2	496,223	180	60	10,719	83	605	0.05	78	1
Uruguay	3.3	68,039	47	77	15,911	60	630	0.05	75	0
Venezuela	27.6	352,143	316	33	13,616	71	1,35	0.11	60	1
World	**~7 billion**		**70,202**		**11,975**		**1,191,707**	**100**		

[a] *Source: www.infoplease.com/ipa/A0004379.html.* [b] *GDP—gross domestic product. Source: http://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal).* [c] *Source: http://en.wikipedia.org/wiki/List_of_countries_by_GDP_(PPP)_per_capita.* [d] *Source: Institute for Scientific Information, National Science Indicators, USA. Base Deluxe–SCI (2009).* [e] *Source: PubMed using the terms pharmacogen* AND country name, http://www.ncbi.nlm.nih.gov/sites/entrez, accessed April 15, 2011*

PGx experts were invited to contribute keynote lectures and to participate in a number of symposia. The conference covered topics related to the design, interpretation, and implications of PGx studies in Latin America, but also took on a more global approach to the growing contribution of genomic information and biotechnological tools to the way medicines are developed, regulated, and prescribed to patients [8].

In accordance with its mission to provide leadership in PGx research and education in Brazil, with impact on population health, Refargen has recently concluded a study of the distribution of PGx polymorphisms in a large, representative cohort of the overall Brazilian population [2,9–11]. The data generated by this study, available online [12], confirmed the substantial PGx diversity among Brazilians, depending on biogeographical ancestry,

TABLE 45.2 Research Areas of Refargen Investigators

Pharmacogenomics
Anticoagulants, antiplatelet agents, and statins
Cancer
Infectious diseases (HIV/AIDs, Hanseniasis, Leishmaniosis)
Kidney transplant
Modulation of premature stop codons
Neurological and psychiatric disorders
Population pharmacogenetics
Population genetics/human genome diversity
Forensic genetics and biological anthopology
Bioinformatics and medical genomics/biological modeling
Clinical and preclinical pharmacology/clinical pharmacokinetics
Geriatrics, atherosclerosis, and dislipidemias
Immunogenetics
Molecular biology of multifactorial diseases
Molecular and genetic hematology
Translation investigation in oncology

Source: www.refargen.org.br/rubrique.php3?id_rubrique=2, accessed May 31, 2013.

self-reported “race/Color”[1] and geographical origin. The following sections review these findings and discuss their PGx implications.

45.2.1 Pharmacogenomic Diversity among Brazilians

The heterogeneity of the Brazilian population, well documented in a wealth of population genetic studies [2,13], is reflected in the distribution of PGx polymorphisms. Admixture of Amerindian, European, and sub-Saharan African parental populations is the major determinant of heterogeneity among Brazilians. Panels of ancestry-informative markers (AIMs) have been used to obtain estimates of the individual proportions of these ancestries in Brazilians. Figure 45.1 shows results obtained with a panel of insertion–deletion polymorphisms developed by Professor Sergio D. J. Pena from the Federal University of Minas Gerais, and validated as AIMs for the Brazilian population [2,14]. The data pertain to healthy, unrelated individuals, self-identified as white (n = 179) or black (n = 177), recruited in Rio de Janeiro and São Paulo, the two largest cities of Brazil. It is evident that the *individual* proportions of Amerindian, European, and African ancestry vary largely and in a *continuous* manner in both Color groups, especially among self-identified blacks for whom estimates of African and of European ancestry range between <0.01–0.99 and <0.01–0.96, respectively.

Another distinct feature of the plots in Figure 45.1 is the extensive overlap of the individual proportions of European and African ancestry in white and black individuals, which supports the notion that among Brazilians, correlation between Color categories and biogeographical ancestry is tenuous [2,15]. Despite the wide range and extensive overlap of individual proportions of European and African ancestry in both the black and the white cohorts in Figure 45.1, the average contribution of European ancestry is considerably higher in white (0.92, 0.86–0.95; median, first to third interquartile range) than black persons (0.34, 0.13–0.69), whereas African ancestry is much larger in self-identified blacks (0.54, 0.22–0.77) than whites (0.03, 0.02–0.06). The average proportion of Amerindian biogeographical ancestry in the white and black cohorts shown in Figure 45.1 is very similar, with median (interquartile range) values of 0.03 (0.02–0.06) and 0.04 (0.02–0.11), respectively.

The data used to construct the plots of Figure 45.1 were obtained from individuals recruited in the Southeast region of Brazil. Extrapolation to the overall Brazilian population must take into account the great variability in the relative proportions of Amerindian, European, and African ancestries among Color categories in different geographical regions [2]. Nevertheless, two major features of Figure 45.1 are common to the vast majority of Brazilians, and impact the distribution of PGx polymorphisms: (1) the European and African components together account for ∼90% of the biogeographical ancestry and (2) the individual proportions of European and African ancestry vary over wide ranges and, most important, as a continuum irrespective of self-reported color and geographical origin. Based on these two features, it is reasonable to anticipate that:

- The greater the difference in frequency of a given PGx polymorphism between European and sub-Saharan

---

[1] This chapter uses the race/Color classification adopted by the Brazilian Census, which relies on self-assessment of skin Color. Five categories are proposed: *branca* (white), *parda* (brown), *preta* (black), *indígena* (Amerindian), and *amarela* (yellow, meaning of Asian extraction). The term *Color* will be capitalized hereafter to call attention to its special meaning in the context of the Brazilian Census classification.

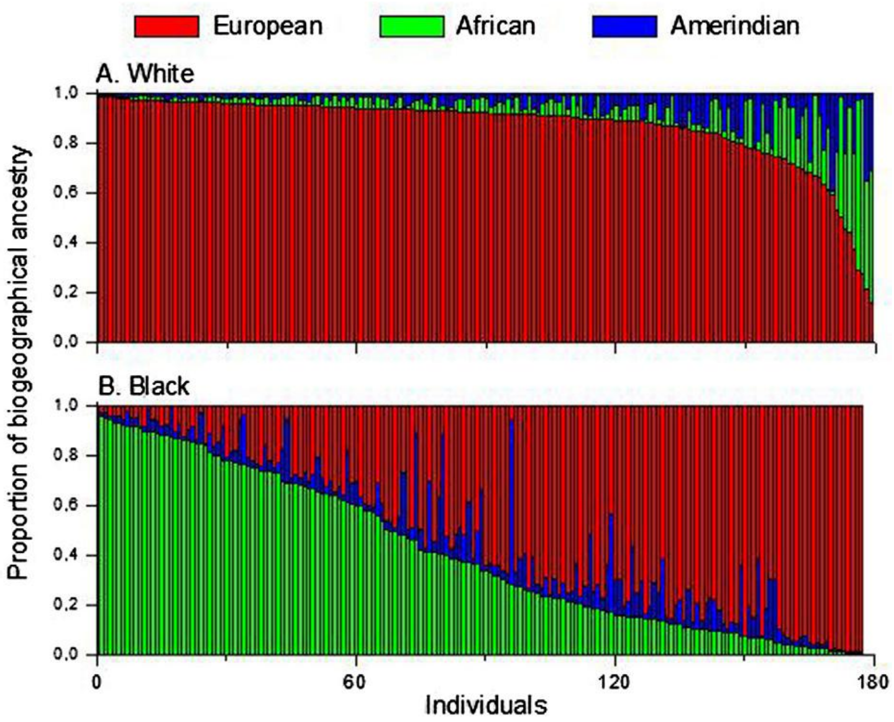


FIGURE 45.1 **Assignment of Brazilian samples to genetic clusters.** The genetic clusters are inferred from a panel of ancestry-informative markers and the program STRUCTURE. Data from unrelated, healthy adult residents in the Southeast region of Brazil, self-identified as white (n = 179) or black (n = 177), according to the color classification adopted by the Brazilian Census. Each individual is represented by a vertical line, which is partitioned in three colored segments corresponding to the estimated membership proportion in the African (blue), Amerindian (green), and European (red) clusters. *Source: Data from Suarez-Kurtz et al.* [9].

African ancestral populations, the more likely the polymorphism frequency will vary within the admixed Brazilian population.

- The difference in frequency of PGx polymorphisms between Europeans and Africans will be attenuated in white versus black Brazilians.
- The frequency of the PGx polymorphisms in Brazilians will vary continuously as a function of the individual proportions of European and/or African ancestry.

These predictions have been verified for a number of polymorphic pharmacogenes, including *CYP3A5* [16], *ABCB1* [11,17], *VKORC1* [9], and the *CYP2C* cluster [10]. The study of *VKORC1* was particularly informative since comparative data were obtained from Portuguese, by far the most important source of European migrants to Brazil, and sub-Saharan Africans from former Portuguese colonies—the origin of enslaved Africans brought to Brazil. The Brazilian cohort consisted of 1037 healthy, unrelated adults, self-categorized as white (n = 342), brown (350) or black (345), recruited in the context of a Refargen collaborative project in the North, Northeast, South and Southeast regions.

We consider this cohort representative of the present-day Brazilian population because 99% of Brazilians self-identify with one of these Color categories and 93% live in the geographical regions included in the study [18]. The African cohort consisted of 216 healthy adults from Cabinda, Angola (n = 73), Maputo, Mozambique (n = 143), and the Portuguese sample comprised 89 healthy adults recruited in Oporto.

The procedures used for allele discrimination of the *VKORC1* 3673G>A (rs9923231, standard coordinate –1639G>A, based on Genbank accession number AY587020), 5808T>G (rs2884737, 497T>G), 6853G>C (rs8050894, 1542G>C), and 9041G>A (rs7294, 3730G>A) single-nucleotide polymorphisms (SNPs) and for inference of their haplotypes were described in the original report [9]. Five haplotypes, namely GTGA, AGCG, GTGG, ATCG, and GTCG, occurred with frequencies >1% and accounted for >99% of the genetic diversity in Brazilians, Portuguese, and Africans (Table 45.3).

The haplotype frequency in white and black Brazilians was "weighted" according to the relative proportions of individuals self-identified with each Color in each geographical region, in the most recent Brazilian Census [19]. To give an example, white individuals in the North, Northeast, Southeast, and South regions represent, respectively, 18.3%, 4.4%, 52.1%, and 25.2% of the sum of white individuals in these four regions [18]. By multiplying these

TABLE 45.3 *VKORC1* Haplotype Frequency in Brazilians, Portuguese, and Sub-Saharan Africans (Angolans and Mozambicans)

Haplotype[a]	White Brazilians (347)[b]	Black Brazilians (353)[b]	Portuguese (89)[b]	Africans (216)[b]
GTGA	0.37 (0.32–0.42)	0.38 (0.33–0.43)	0.28 (0.21–0.34)	0.39 (0.34–0.44)
GTGG	0.22 (0.18–0.27)	0.26 (0.21–0.31)	0.26 (0.20–0.33)	0.36 (0.32–0.41)
AGCG	0.25 (0.20–0.30)	0.12 (0.09–0.16)	0.29 (0.22–0.36)	0.01 (<0.01–0.02)
ATCG	0.12 (0.08–0.16)	0.12 (0.09–0.16)	0.12 (0.8–0.18)	0.02 (<0.01–0.03)
GTCG	0.03 (0.02–0.05)	0.12 (0.09–0.16)	0.03 (<0.01–0.05)	0.21 (0.17–0.25)

[a] *VKORC1 haplotypes comprising rs9923231, rs2884737, rs8050894, and rs7294.* [b] *Number of individuals. Haplotype frequency expressed as mean (CI95%).*

Note: Sub-Saharan Africans include Angolans and Mozambicans.

percentages by the haplotype frequency in whites in the corresponding region, the weighted frequency for white Brazilians is obtained. A similar procedure was used to estimate the haplotype frequency in black Brazilians. The distribution of the *VKORC1* haplotypes did not differ between Angolans and Mozambicans ($p = 0.57$, χ^2 test) and the two datasets were combined in Table 45.3.

The distribution of the *VKORC1* haplotypes differed significantly ($p < 0.0001$) among the four cohorts shown in Table 45.3. Pairwise comparisons disclosed significant differences ($p < 0.0001$) in haplotype distribution between Portuguese versus Africans and white versus black Brazilians. However, as predicted by the European/African admixture of the Brazilian population (see preceding), the difference in frequency of each haplotype was markedly attenuated in white versus black Brazilians, compared to Portuguese versus Africans (Figure 45.2(a)).

In Figure 45.2(b), we contrast the differences in frequency of the *VKORC1* haplotypes by means of Wright's F-statistics (F_{ST}), a simple and efficient parameter to assess population diversity [20]. F_{ST} values are calculated using the expression

$$F_{ST} = \frac{(p_1 - p_2)^2}{(p_1 + p_2)(2 - p_1 - p_2)}$$

where p_1 and p_2 denote the frequency of a given allele or haplotype in population 1 and population 2, respectively. F_{ST} has a lower bound of 0, when the allele frequency is the same in the two populations, and an upper value of 1, when the allele frequency is 0 in one population and 1 in the other.

The extent of PGx divergence between any two populations is denoted according to Wright's qualitative guidelines [21]: F_{ST} values <0.05 represent low genetic divergence, between 0.05–0.15 represent moderate divergence, between 0.15 and 0.25 represent large divergence, and values >0.25, very large divergence. Accordingly, haplotypes AGCG and GTCG display large and moderate genetic divergence, respectively, in Portuguese versus Angolans/Mozambicans, but only low divergence in white versus black Brazilians, as a predicted consequence of Brazilian population admixture.

The F_{ST} analyses were extended to 44 individual polymorphisms in 12 important pharmacogenes affecting drug metabolism, transport, and effect, which had been genotyped in the Refargen cohort [19]. Pairwise comparisons between white, brown, and black groups in the entire cohort revealed low genetic divergence in general, with mean F_{ST} values of 0.005 (SD = 0.006), 0.013 (0.017), and 0.004 (0.005) for white versus brown, white versus black and brown versus black, respectively.

Only three SNPs, namely *CYP3A5*3* and the linked *SLCO1B3* 334T>C and 699G>A transitions exceeded the F_{ST} threshold for moderate genetic divergence (0.05) in white versus black Brazilians (Figure 45.3). For comparison, this figure shows the F_{ST}'s for the same polymorphisms in the HapMap CEU versus YRI groups, taken as proxies of European and sub-Saharan African ancestral populations. The F_{ST} values varied over a wide range (0.000–0.624) with a mean of 0.087 (SD = 0.118); 9, 4, and 4 SNPs displayed moderate, large, and very large divergence, respectively. The attenuation of pharmacogenetic divergence between white and black Brazilians compared to the HapMap populations is evident.

We suggested above that, as a predictable consequence of admixture, the frequency of PGx polymorphisms will vary continuously among Brazilians, as a function of the individual proportions of European and/or African ancestry. Logistic modeling using maximum likelihood estimation has been used by Refargen members to explore the association between PGx polymorphisms and the estimated individual biogeographical ancestry. A consistent finding in these analyses [9,10,16,17,22,23] is that the frequency

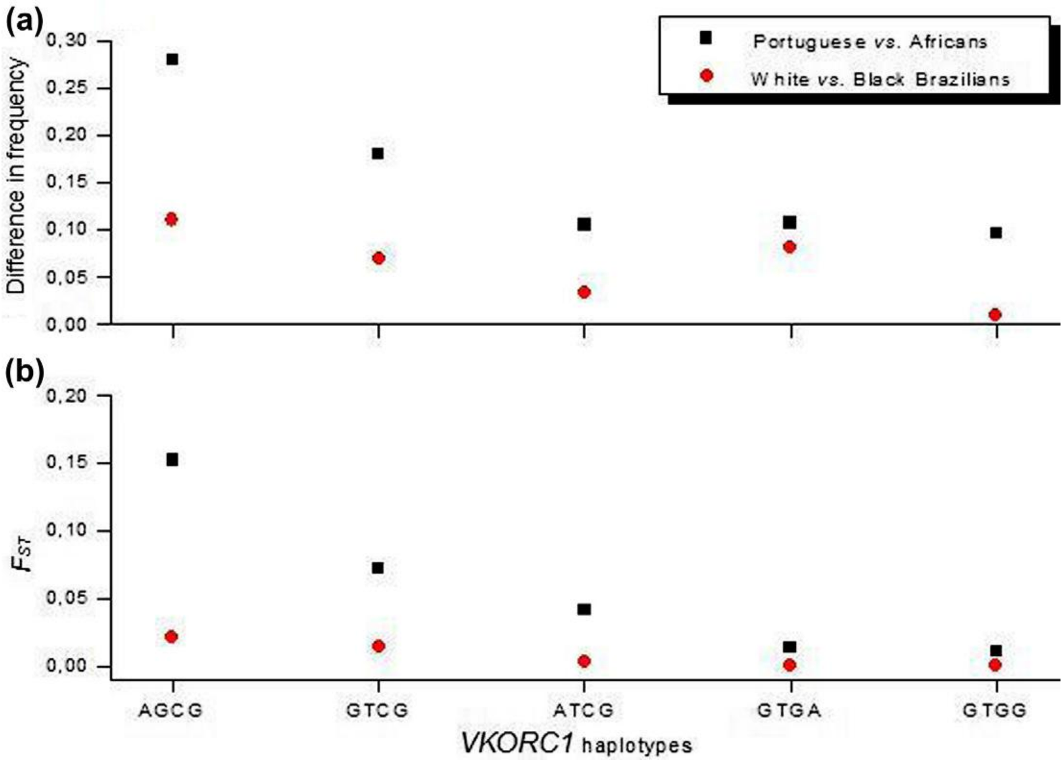


FIGURE 45.2 **Distribution of *VKORC1* haplotypes.** White (n = 347) and black (n = 353) Brazilians, Portuguese (n = 89), and Sub-Saharan Africans from Angola (n = 73) and Mozambique (n = 213). The haplotypes, comprising the rs9923231, rs2884737, rs8050894, and rs7294 SNPs, were inferred using the *Haplostats* software [61]. (a) Difference in haplotype frequency between white and black Brazilians and between Portuguese and Africans. (b) F_{ST} values for white versus black Brazilians and Portuguese and Africans. *Source: VKORC1 haplotype data from Suarez-Kurtz et al.* [9].

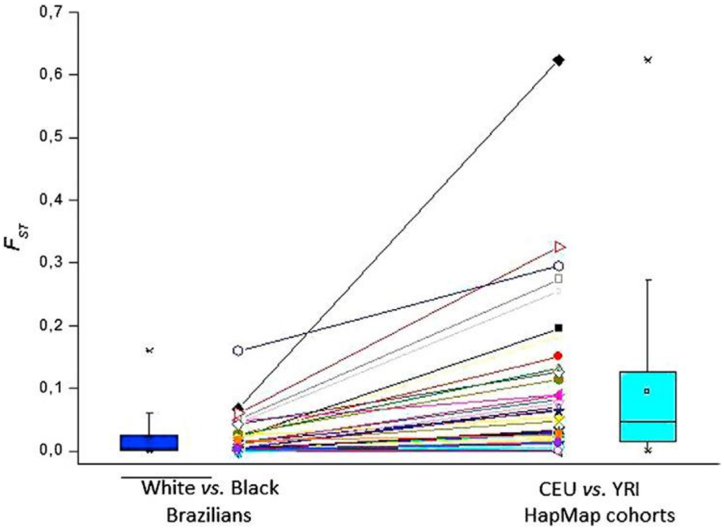


FIGURE 45.3 **Allele-specific F_{ST} values.** F_{ST} values are for 40 PGx polymorphisms in HapMap CEU versus YRI and in white versus black Brazilians. The lines connect the F_{ST} values for each polymorphism in the two datasets and the box plots at the left and right summarize the ensemble of the data for each set. *Source: Data for Brazilians from Suarez-Kurtz et al.* [19] *and for CEU and YRI from the International HapMap project* [81].

distribution of PGx polymorphisms among Brazilians is best fit by continuous functions of the individual proportions of African and European ancestry. This is illustrated in Figure 45.4 with data for the *VKORC1* 3673G>A genotype, which is of major clinical relevance in relation to warfarin dosing and response in the Brazilian population (see below). The plot shows that the odds of having the heterozygous, and to a lesser extent, the homozygous variant genotype

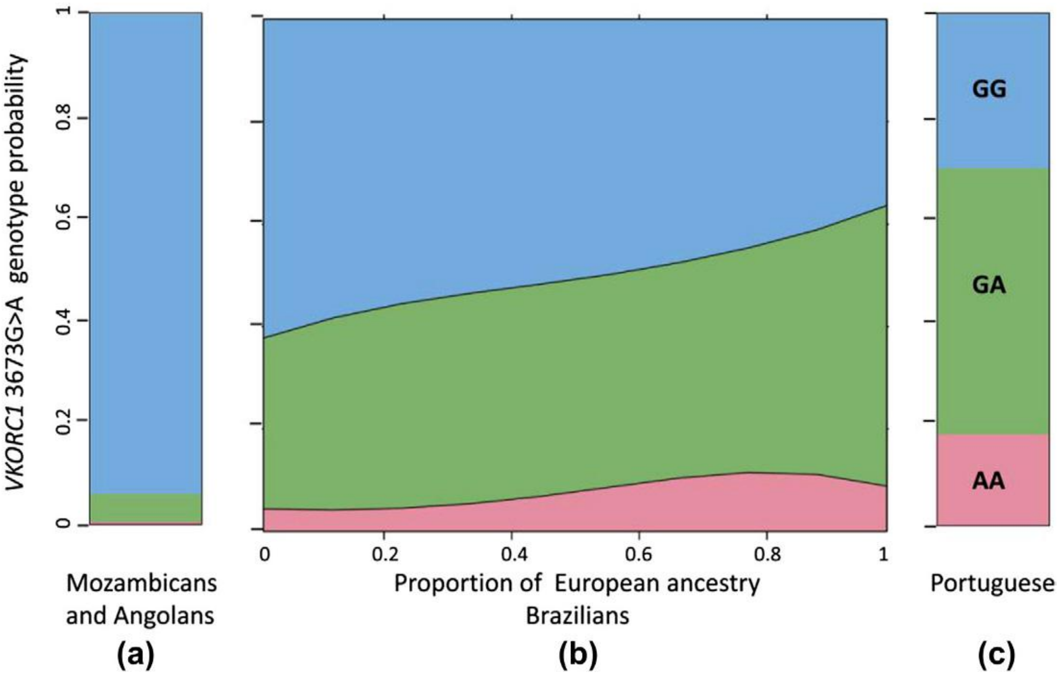


FIGURE 45.4 Effect display for the distribution of *VKORC1* 3673G>A genotypes in a logit model. The genotypes are fit to the data for African ancestry in 965 Brazilians (panel B). For comparison the frequency of each genotype in a cohort of Angolans (n = 73) and Mozambicans (143) (panel A) and in a Portuguese cohort (89) (panel C) are also shown. The individual proportion of African ancestry in Brazilians is shown on the x-axis. The y-axis represents the genotype probability for Brazilians and the observed genotype frequency for the African and Portuguese cohorts. *Source: Data from Suarez-Kurtz et al.* [9].

at the *VKORC1* 3673G>A locus increase progressively as the individual proportion of European ancestry increases in the overall cohort, irrespective of self-identified Color. For comparison, we also display in Figure 45.4 the distribution of *VKORC1* 3673G>A genotypes among Portuguese and sub-Saharan Africans, from Angola and Mozambique.

45.2.2 PGx Clinical Trials in Brazilians

The heterogeneity of the Brazilian population renders inappropriate the extrapolation of PGx data derived from well-defined ethnic groups. Only recently, recognition of this fact translated into PGx research on the clinical response to prescribed drugs in Brazilians, as illustrated in the following selected examples from studies carried out by Refargen groups:

- Rosario Hirata studied the PGx of HMG-CoA reductase inhibitors [24–26] and clopidogrel [27]
- Rosane Vianna-Jorge examined the impact of polymorphisms in the cyclooxygenase-2 gene and breast cancer risk and prognosis [28–30]
- Vanessa Mattevi investigated the *UGT1A1*28* variant allele as a predictor of severe hyperbilirubinemia in HIV-infected patients receiving highly active antiretroviral therapy (HAART) [31]
- Adalberto Santos and Ândrea K. Ribeiro dos Santos, in collaboration with Mara Hutz, explored the association between N-acetyltransferase 2 (NAT2) haplotypes and the hepatotoxicity induced by antituberculosis drugs in Southern and Northern Brazilians, respectively [32,33]

Our own studies covered the PGx of antipsychotics, HAART, immunosuppressants, L-thyroxine, levodopa, methylphenidate, nonsteroidal antiinflammatory agents (NSAIDs), statins, and warfarin. In the following sections, we review results from some of these studies.

45.2.2.1 Pharmacokinetics and Pharmacodynamics of NSAIDs

The implementation of regulatory legislation for generic drugs in Brazil during the late 1990s provided an impetus for local bioequivalence studies. The relatively large number of samples collected from healthy subjects under the rigorously controlled conditions of bioequivalence trials offer a favorable experimental setting for assessing the influence of PGx polymorphisms on pharmacokinetic parameters. We explored this opportunity to examine the impact of the defective *CYP2C9*2* and *CYP2C9*3* alleles on the pharmacokinetics and pharmacodynamics of NSAIDs, tenoxicam, and piroxicam [34,35]. Serial blood samples were collected for 240 hours following single oral doses (20 mg) of piroxicam (n = 35) or tenoxicam (n = 24); drug concentrations in plasma were measured by validated HPLC methods and

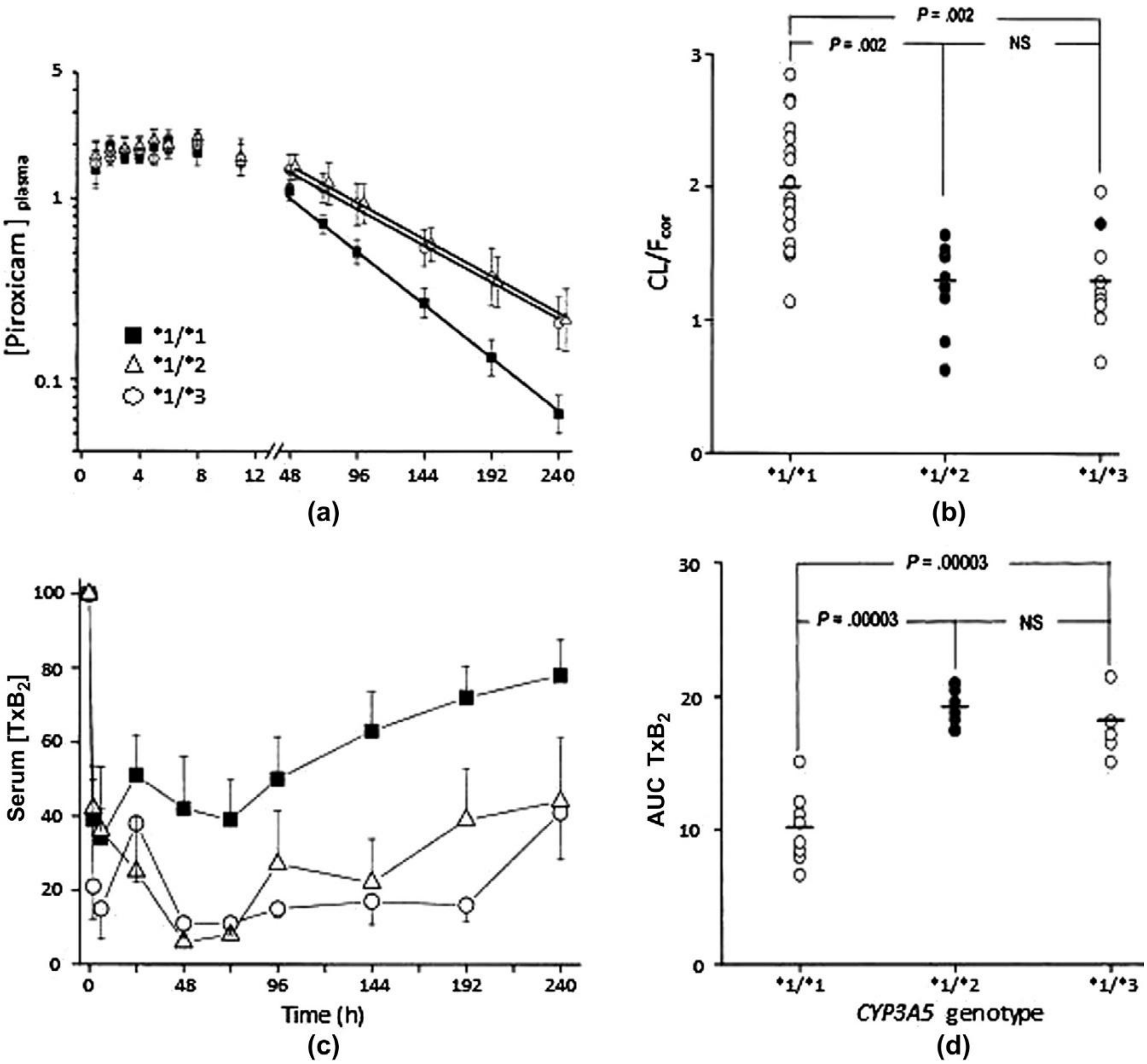


FIGURE 45.5 Influence of *CYP2C9* polymorphisms on the pharmacokinetics and pharmacodynamics of piroxicam. Data from 35 healthy subjects, expressing different *CYP2C9* genotypes (*CYP2C9*1/*1*, n = 19; *CYP2C9*1/*2*, n = 9; *CYP2C9*1/*3*, n = 9), after a single oral dose of piroxicam (20 mg) administered at time zero. (a) Piroxicam plasma concentration (μg/mL^{-1})-time curves; the lines are the best linear fit for the data points on the elimination phase (48–240 h). (b) Individual values of the oral clearance of piroxicam corrected for body weight (CL/F$_{cor}$, expressed as mL h^{-1} kg^{-1}. (c) Effect of piroxicam on the serum concentration time curves of TxB$_2$. Data are normalized to the baseline serum TxB$_2$ concentration and presented as mean ± SEM. (d) Individual values of the area above the time–effect curve from 0 to 240 h (TxB$_2$ AUC). To construct the plots in (d), the serum TxB$_2$ concentration at each time point was normalized to the respective baseline TxB$_2$ concentration, which was taken as 1 and the area above the time–effect curve from 0 to 240 h was integrated. The data are presented as TxB$_2$ h^{-1}). In (b) and (d), the horizontal bars are the median for each *CYP2C9* genotype, *P* values are for pairwise comparisons between *CYP2C9*1/*1* and either *CYP2C9*1/*2* or *CYP2C9*1/*3*, using the Student-Newman-Keuls test and individuals having one or two copies of the *CYP2C8*3* allele are identified by gray and black circles, respectively. NS = nonsignificant. *Source: Data from Perini et al.* [35].

noncompartmental pharmacokinetic analyses were applied to data analyses. Results obtained for both NSAIDS were similar and are illustrated here with data for piroxican.

Figure 45.5 shows the plasma concentration–time curves according to the *CYP2C9* genotypes. The area under the plasma concentration-time curve (AUC) was significantly greater in *CYP2C9*1/*2* or **1/*3* heterozygous, as compared to wild-type homozygous subjects (*CYP2C9*1/*1*). This effect was associated with and explained by the significantly lower oral drug clearance in heterozygous carriers of the *CYP2C9*2* or **3* alleles. There was no significant difference in the AUC or drug clearance between

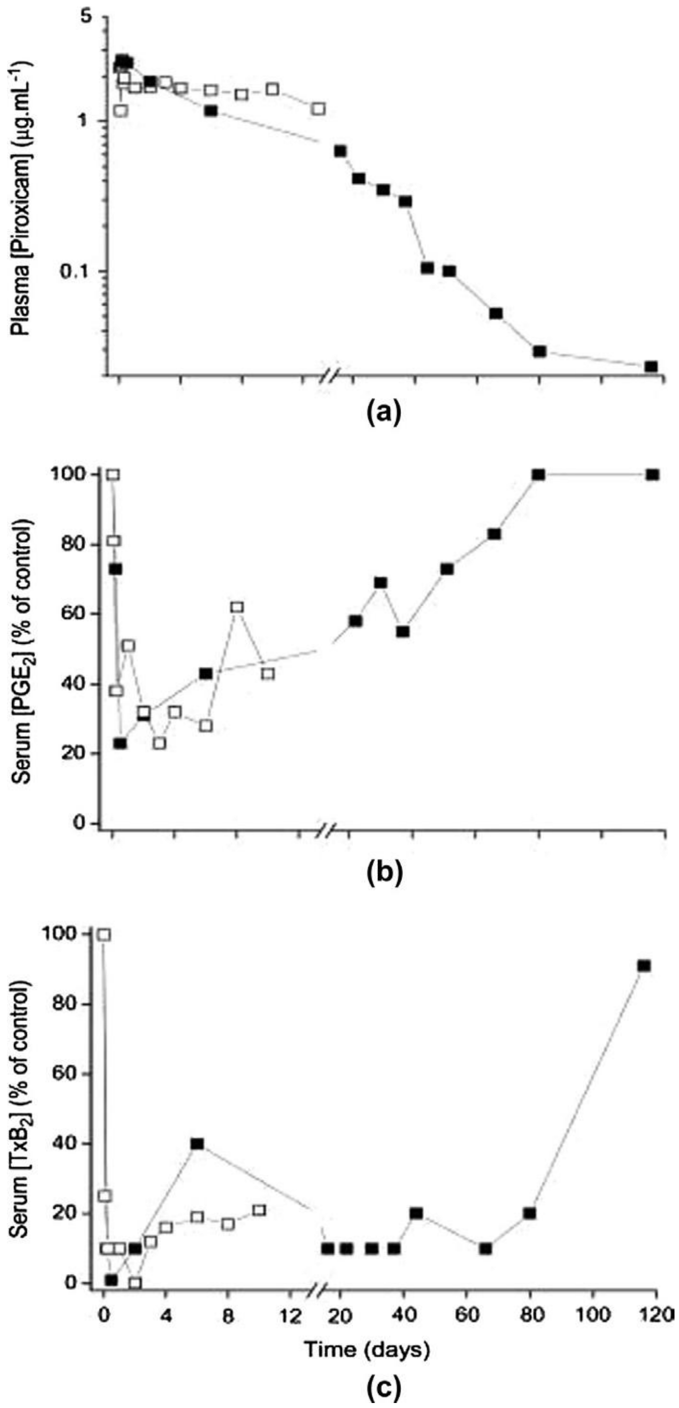


FIGURE 45.6 **Pharmacokinetic and pharmacodynamic data from a subject expressing the *CYP2C9*3/*3* genotype.** The data shown are for a period following administration of single doses of piroxicam (20 mg po) on two different occasions separated by six months. Blood samples were collected at various time intervals for 312 h and for 2784 h following administration of the first (open squares) and the second dose (solid squares), respectively. (a) Plasma concentration–time curves of piroxicam. (b, c) Serum concentration–time curves of PGE$_2$ and TxB$_2$, respectively; the data are normalized to the baseline serum PGE$_2$ and TxB$_2$ concentrations, measured 24 h (first dose) and 24–72 h (second dose) prior to each drug administration. *Source: Reproduced from Perini and Suarez-Kurtz* [36].

the *CYP2C9*1/*2* and **1/*3* heterozygous individuals. One individual in the piroxicam study had the *CYP2C9*3/*3* genotype, which is extremely rare in Brazilians [10].

The pharmacokinetic and pharmacodynamic consequences of this double-mutant genotype following a single dose of piroxicam were remarkable (Figure 45.6): (1) the

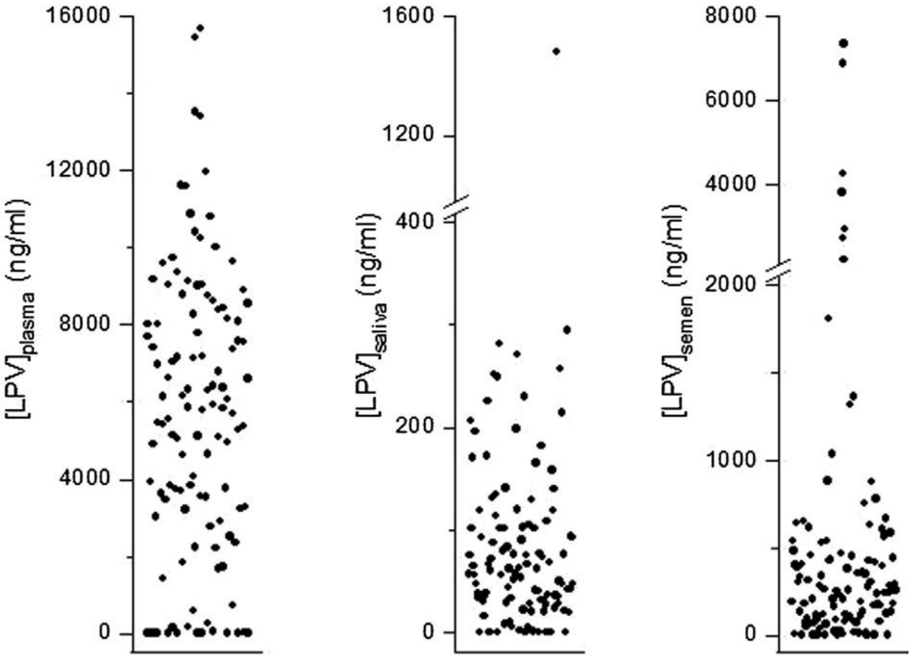


FIGURE 45.7 **Scatter plot of the distribution of trough (predose) concentrations of lopinavir.** Shown are the blood plasma (left panels), saliva (center), and semen (right) among 99 HIV-infected subjects. The individuals were under stable HAART treatment, including two daily doses of Kaletra® (400 mg lopinavir and 100 mg ritonavir). *Source: Data from Estrela et al.* [38].

half-life of elimination was extended to ~500 hours, the activity of cyclo-oxygenase 1 (COX-1), measured by the *ex vivo* production of thromboxane B2 (TxB_2), was maintained below 20% of its control value for more than two months; and (2) COX-2 activity, measured by *ex vivo* production of PGE_2, was reduced to <50% of its control value for more than one month! Nevertheless, no adverse effects were observed. These data led us to speculate on the potential advantages conferred by the PGx polymorphisms associated with impaired drug metabolism for long-term strategies in disease prevention and treatment [36].

45.2.2.2 Pharmacokinetics of Antiretrovirals

The Brazilian public health system pioneered, in the early 1990s, the universal access to HIV medications free of charge to all citizens who need it [37]. Ritonavir-boosted lopinavir formulations are frequently included in HAART regimens for HIV infection. Both these HIV protease inhibitors are substrates for CYP3A drug-metabolizing enzymes and the drug transporters ABCB1 (P-glycoprotein) and SLCO1B1. We studied the impact of polymorphisms in the *CYP3A5*, *SLCO1B1*, and *ABCB1* genes on the pharmacokinetics of lopinavir and ritonavir in 115 HIV-infected men under stable treatment with Kaletra® at the Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro. The study cohort was genotyped for *ABCB1* 1236C>T, 2677G>T/A and 3435C>T, *SLCO1B1* 521T>C, 388A>G, and 463C>A and CYP3A5*3 and *6 polymorphisms. The trough (predose) concentrations of lopinavir and ritonavir in blood plasma, semen, and saliva were measured and validated using LC/MS/MS methods.

Large interindividual variability in the concentrations of both protease inhibitors in the three matrices were observed, as illustrated in Figure 45.7 with data for lopinavir. Univariate and multivariate regression analysis revealed no influence of *ABCB1* genotypes or haplotypes on the concentrations of lopinavir and ritonavir in plasma, semen, and saliva [38].

We took advantage of the availability of plasma concentration data for lopinavir and ritonavir, from this cohort of HIV-infected men under stable HAART to examine the influence of CYP3A5 and SLCO1B1 polymorphisms on the pharmacokinetics of these protease inhibitors. The goal of the *CYP3A5* study [39] was to verify whether the reported impact of the *CYP3A5* genotype on the disposition of the HIV protease inhibitor (saquinavir) administered as a single dose and with no ritonavir boosting to Tanzanian healthy individuals [40] prevailed under a natural clinical setting.

The frequencies of *CYP3A5*3* and *CYP3A5*6* in the study sample were 0.67 (95% CI 0.61–0.74) and 0.05 (0.02–0.08), respectively. Assuming that alleles *CYP3A5*3* and **6* do not occur in the same DNA strand, the study participants were genotyped as **1/*1* (n = 11), **1/*3* (32), **1/*6* (1), **3/*3* (47), **3/*6* (6), and **6/*6* (1). The distribution of the trough concentrations of lopinavir and ritonavir in plasma among individuals having two (n = 11), one (33) or zero (54) *CYP3A5* functional alleles is plotted in Figure 45.8.

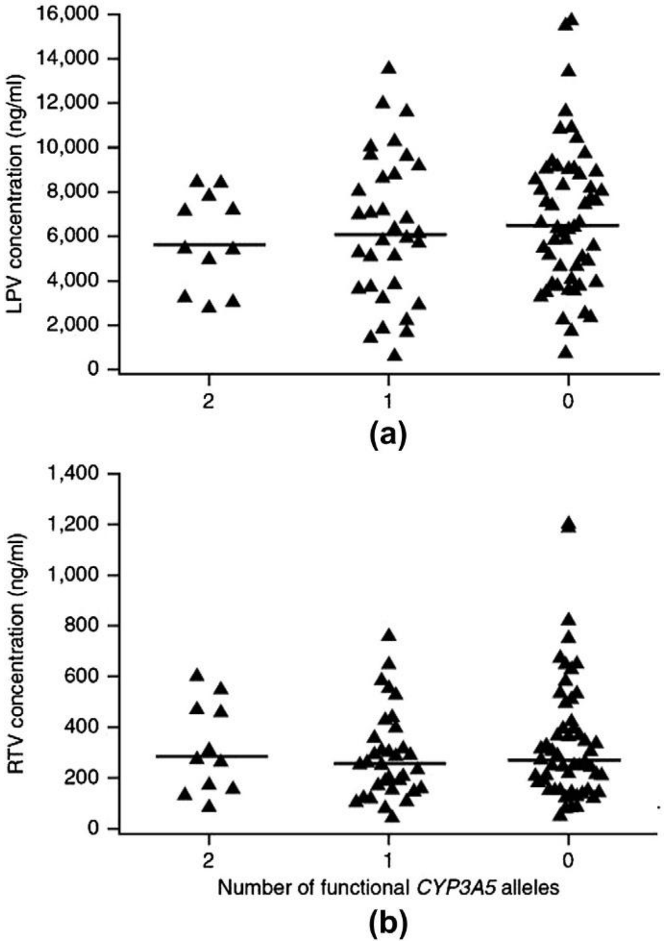


FIGURE 45.8 **Distribution of the trough concentrations of lopinavir (a) and ritonavir (b).** Shown are plasma among HIV-infected subjects with 2 (n = 11), 1 (n = 33), or 0 (n = 54) *CYP3A5* functional alleles (p = 0.65 for lopinavir and *P* = 0.78 for ritonavir; Kruskal-Wallis test). The lines denote the median values for each genotype. *Source: Reproduced from Estrela et al.* [39].

The overlapping of the plasma trough concentrations of each protease inhibitor among the three groups is evident and no statistically significant intergroup difference was detected for either lopinavir or ritonavir (Kruskal-Wallis test). We ascribed the apparent discrepancy between our results [39] and those of Josephson et al. [40] to ritonavir boosting, which was thought to occlude any pharmacokinetic consequences of functional polymorphisms in the *CYP3A5* gene. Together, the two studies highlight complementary steps towards the implementation of PGx in clinical practice: namely, the discovery of association(s) between genetic polymorphisms and pharmacological traits under specific experimental conditions, followed by examination of the impact of the association(s) in the targeted patient populations, studied in the relevant clinical setting.

HIV protease inhibitors are substrates of the SLCO1B1/OATP1 drug transporter, and a SNP (521T>C) in the SLCO1B1 gene was associated with plasma levels of lopinavir in HIV-infected individuals [41]. We examined this association in our cohort of HIV-infected men under stable HAART, and extended our analyses in two directions: (1) by examining two other SLCO1B1 SNPs (388A>G, 463C>A) and (2) by exploring the association of the SLCO1B1 polymorphisms with ritonavir plasma concentrations [42].

We observed no influence of the 388A>G and 463C>A SNPs on lopinavir or ritonavir plasma concentrations, but confirmed that carriers of the variant 521C had significantly higher lopinavir (but not ritonavir) plasma concentrations relative to the wild-type TT genotype. Reduced OATP-mediated uptake of lopinavir by the hepatocytes allele was thought to account for the increased plasma levels of lopinavir in carriers of SLCO1B1 521C. The clinical utility of this observation is uncertain, in view of both the extensive overlap of the trough concentrations of lopinavir in the plasma across the three 521T>C genotypes and the low prevalence (<3%) of the homozygous variant genotype (521CC) in Brazilians [43].

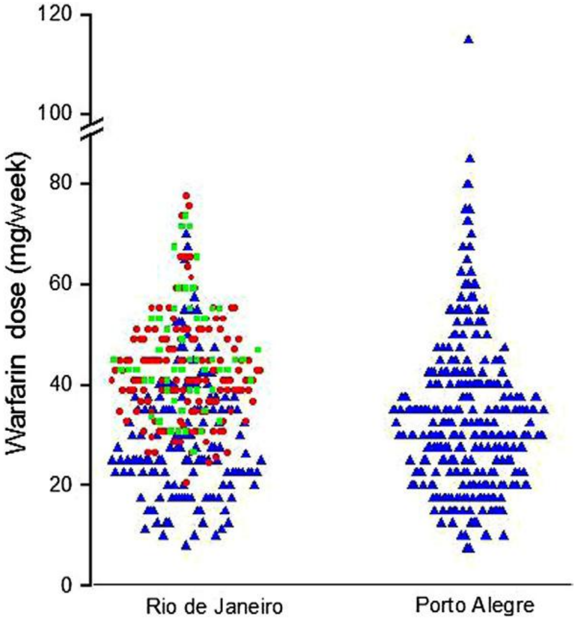


FIGURE 45.9 **Distribution of stable warfarin doses (mg week^{-1}) in two cohorts of Brazilian patients.** The Rio de Janeiro cohort (n = 390) included individuals who self-identified as white (n = 96, blue triangles) brown (118, red circles), or black (76, green squares). All Porto Alegre patients (n = 302, blue triangles) were white of exclusively European descent. *Source: Data from Suarez-Kurtz et al.* [47] *and Botton et al.* [4].

45.2.2.3 Warfarin-Dosing Algorithms for Brazilians

Warfarin combines several characteristics that make it a case study for personalized drug therapy. Several PGx algorithms have been described for the prospective estimation of individual warfarin dose requirements, but their performance varies considerably across populations [44]. This discrepancy is explained, in part, by the frequency distribution of polymorphisms in genes that influence warfarin pharmacokinetics (*CYP2C9*) and pharmacodynamics (*VKORC1, CYP4F2*). Our two groups examined the influence of polymorphisms in these and other pharmacogenes, as well as demographic and clinical variables on the interindividual variability in warfarin dose requirements for stable anticoagulation in Brazilian patients (Table 45.4). We then applied a two-step (univariate followed by multivariate) multivariate regression modeling to develop dosing algorithms for prediction of the appropriate warfarin dose for each patient [45–49].

The cohort studied by Suarez-Kurtz and colleagues consisted of 390 adult, ambulatory patients from a public hospital in Rio de Janeiro, self-identified as white (n = 196), brown (118), or black (76), under stable warfarin therapy. The individual warfarin dose requirement ranged from 3–78 mg/week (i.e., a 26-fold variation) and differed significantly ($P = 0.0004$, ANOVA) among white (28.9 ± 12.3, mean ± SD), brown (32.9 ± 12.4), and black patients (35.3 ± 14.6) (Figure 45.9). In the two-step regression analyses, the *VKORC1* 3673G>A genotype proved to be the most important predictor of warfarin dose, followed by the number of copies of variant *CYP2C9* alleles and cotreatment with amiodarone. Body weight, therapeutic indication for warfarin, age, and cotreatmtent with simvastatin were the other covariates associated with warfarin dose requirement.

A dosing algorithm with these covariates explained 51% ($R^2 = 0.51$) of the variability in warfarin dose requirement in the overall cohort and, most importantly, performed equally well in self-reported white and black individuals. This contrasts markedly with the considerably poorer performance of warfarin algorithms in African Americans versus white Americans (e.g., $R^2 = 0.51$ vs. 0.31) [50], which is not reversed by inclusion of a "race" term in the algorithms [50,51].

This discrepancy may be explained by greater admixture among Brazilians compared to Americans, in particular black Brazilians versus African Americans [1]. The *CYP4F2* rs2108622 (V33M) SNP had a marginal effect on the predictive power of the warfarin-dosing algorithm in this cohort [46], whereas the −401G>T and −402G>A SNPs in the coagulation Factor VII gene had no effect [52]. Subsequently, we showed that inclusion of a term relating the prescribed warfarin dose to the corresponding international normalized ratio (INR) value (INR/weekly dose) increased the predictive power of our warfarin-dosing algorithm to $R^2 = 0.61$ [47], which remains the highest among algorithms developed for admixed cohorts worldwide.

Data from the cohort studied by Suarez-Kurtz were included in the analyses carried out by the International Warfarin Pharmacogenetics Consortium [53], which led to the development of a "global" warfarin algorithm with "race" as a covariate, distinguishing white, Asian, black or African American, and mixed individuals. Despite the race

TABLE 45.4 Warfarin Dosing Algorithms in Brazilian Cohorts

Model parameters	Description	Perini et al. [44]	Suarez-Kurtz et al [46]	Botton et al [47]	
				Model 1	Model 2
Age		▨	▨	▨	▨
Weight		▨	▨	▨	▨
Co-morbidities	Heart valve prosthesis	▨	▨		
	Thromboembolic disease	▨	▨		
Co-medication	Amiodarone	▨	▨	▨	▨
	Simvastatin			▨	▨
	Carbamazepine			▨	▨
	b-adrenoceptor blockers -			▨	▨
	Amlodipine			▨	▨
	Diuretics			▨	▨
CYP2C9 polymorphisms	number of *2, *3 and *5 alleles	▨	▨	▨	▨
VKORC1 polymorphisms	VKORC1 3673G>A genotype	▨	▨	▨	▨
	2 copies of GCG VKORC1 haplotype			▨	▨
	1 copy of GCA VKORC1 haplotype			▨	▨
	2 copies of GCA VKORC1 haplotype			▨	
	1 copy of ATG VKORC1 haplotype			▨	
	2 copies of ATG VKORC1 haplotype			▨	
CYP4F2 polymorphism	rs2108622C>T, V33M			▨	
INR/dose			▨		
Coefficient of determination	R^2	0.51	0.60	0.63	0.58
MAD[a]	mg/week	6.9	6.5	6.9	7.1

[a] *MAD, mean absolute difference between observed and predicted warfarin weekly dose.*

The hatched fields indicate the covariates included in the respective algorithms.

term, the performance of the IWPC algorithm was considerably better in white ($R^2 = 0.45$) than in Asian ($R^2 = 0.32$) or black patients ($R^2 = 0.29$), which supports the notion that the complexity of human diversity cannot be adjusted simply by a numerical coefficient.

A second, larger study of warfarin PGx in Brazilians was carried out by Mara Hutz and colleagues on 302 patients of exclusively European ancestry recruited in Porto Alegre, the largest the city in the South region of Brazil [48]. Extensive intraindividual variation in stable warfarin dose was observed (see Figure 45.9), and two dosing algorithms, comprising 12 and 18 PGx (polymorphisms in *VKORC1*, *CYP2C9*, and *CYP4F2*), clinical, demographic, and medication covariates, were derived.

The predictive power of these algorithms, $R^2 = 0.58$ and 0.63, respectively, are among the highest reported in the literature. However, when applied to the admixed Brazilian cohort from Rio de Janeiro, these algorithms did not reproduce their original performance: the R^2 values were 0.44 and 0.46, respectively, in the overall cohort, increasing to 0.49–0.50 in self-reported whites and declining to 0.37–0.40 in brown and black patients [49]. By the same token, the algorithm derived for the Rio de Janeiro cohort [45] had a comparatively inferior performance in the patients from Porto Alegre [48]. The poorer performance of the algorithms developed in one cohort when applied to the other may be in part a result of differences in biogeographical ancestry between the two study cohorts, exclusively European in Porto Alegre study versus admixed European/African/Amerindian in Rio de Janeiro.

Although these results seem consistent with the notion that warfarin-dosing algorithms "need to be population specific" [48], it is noteworthy that five algorithms derived from Caucasian patients performed poorly when tested in the Porto Alegre cohort of Brazilians of exclusively European ancestry [48]. Does that imply a "cohort specificity" *within* populations of European descent? Assuming that the frequency distribution of the relevant PGx polymorphisms is relatively constant among European-derived populations, does the observed discord in performance of warfarin algorithms reflect a greater than acknowledged contribution of nongenetic factors (e.g., diet, concomitant medication, concurrent diseases, compliance to warfarin prescription)? These are open questions at the present time that deserve further investigation.

45.2.2.4 Pharmacogenetics of L-Thyroxine in Differentiated Thyroid Cancer

Interindividual variation in T_4 doses required for adequate TSH suppression following thyroid ablation is considerable and involves multiple factors, including diet; comedication; comorbidities; and, possibly, genetic modulation of T_4 pharmacokinetics. T_4 has little biological activity per se, and is converted to the active form, triiodothyronine mainly in liver and kidney. Another major metabolic pathway for T_4 in humans is glucuronidation in the liver, mediated by UDP-glucuronosyl transferases (UGTs), mainly the UGT1A1 and UGT1A3 isoforms. We investigated the association between the required T_4 dose and a common *UGT1A1* polymorphism, namely *UGT1A1* –53$(TA)_n$, which consists of a variable dinucleotide repeat within the promoter TATA element, where $n = 5–8$; the longer the repeat allele, the lower the corresponding *UGT1A1* gene expression [54]. Patients were distributed in three groups according to the number (0, 1, or 2) of the low-expression $(TA)_7$ and $(TA)_8$ alleles they carried. A significant trend for decreasing T_4 dose with increasing number of copies of $(TA)_7$ and $(TA)_8$ and a significant difference in T_4 dose across the *UGT1A1* –53$(TA)_n$ genotypes were observed [55].

These results provided the first evidence for association between *UGT1A1*-53$(TA)_n$ and T_4 dose requirement for TSH suppression in a natural clinical setting. Of note, this study reflects real-life community prescribing and dispensing of T_4 in the context of public hospitals in a developing Latin American country. The relatively small (<20%), although statistically significant, difference in T_4 doses across the three *UGT1A1*-53$(TA)_n$ groups, combined with the considerable overlap in these doses, must be taken into consideration when assessing the clinical usefulness of the present results, which should be prospectively tested in an appropriately large cohort, and possibly extended to include polymorphisms in *UGT1A3* and *UGT1A9*.

45.3 PGX OF IMMUNOSUPPRESSANTS IN RENAL TRANSPLANT PATIENTS

The calcineurin inhibitors, cyclosporine A (CSA) and tacrolimus (TAC) are widely used to prevent allograft rejection after solid organ transplantation. Both these drugs display a narrow therapeutic index and high interindividual pharmacokinetic variability, such that monitoring their blood levels is required to improve efficacy and reduce toxicity. Achieving target therapeutic trough (predosing) blood levels (C_0) is of critical importance for CSA- and TAC-induced immunosuppression, especially during the initial period after transplantation when the risk of organ rejection is highest [56]. Multiple factors modulate the variability in C_0 of CSA and TAC, including polymorphisms in genes encoding for drug-metabolizing enzymes and transporters. CSA and TAC are metabolized by the CYP3A subfamily and are substrates of P-glycoprotein, a product of the *ABCB1* gene.

Polymorphisms in *CYP3A4, CYP3A5,* and *ABCB1* have been investigated as modulators of the pharmacokinetics and clinical effects of CSA and TAC in European, North American, and Asian populations with controversial results [57–60]. The extensive variation in worldwide frequency distribution of *CYP3A* and *ABCB1* polymorphisms is a caveat against extrapolation of these data to Brazilians. Recognition of this fact, and the availability of a large cohort of renal transplant patients from a reference center in Brazil,

TABLE 45.5 Influence of *CYP3A5* and *ABCB1* Haplotypes on CSA and TAC C_0/dose (ng·ml^{-1}/mg)

Haplotype	1 Week	1 Month	2 Month	3 Month	1 Week	1 Month	2 Month	3 Month
CYP3A5[a]	**CSA (ng·ml^{-1}/mg)**				**TAC (ng·ml^{-1}/mg)**			
0	0.79 ± 0.35 (5)	0.66 ± 0.13 (5)	0.86 ± 0.64 (5)	0.84 ± 0.34 (5)	0.91 ± 0.36 (9)	1.46 ± 0.75 (9)	1.42 ± 0.91 (8)	1.41 ± 0.44 (9)
1	1.14 ± 0.51 (57)	1.43 ± 0.74 (58)	1.46 ± 0.92 (56)	1.50 ± 0.64 (53)	0.87 ± 0.40 (48)	1.26 ± 0.71 (46)	1.15 ± 0.36 (41)	1.21 ± 0.40 (40)
2	2.10 ± 1.17 (89)	2.69 ± 1.53 (91)	2.88 ± 1.51 (91)	3.11 ± 1.78 (83)	1.00 ± 0.96 (93)	1.37 ± 0.65 (86)	1.31 ± 0.53 (80)	1.40 ± 0.97 (69)
ABCB1[b]	**CSA (ng·ml^{-1}/mg)**				**TAC (ng·ml^{-1}/mg)**			
0	1.94 ± 1.09 (27)	2.25 ± 1.22 (27)	2.54 ± 1.78 (28)	2.72 ± 1.22 (28)	0.94 ± 0.42 (36)	1.41 ± 0.84 (34)	1.54 ± 0.72 (29)	1.63 ± 0.70 (29)
1	1.61 ± 0.92 (88)	2.09 ± 1.40 (91)	2.35 ± 1.47 (90)	2.33 ± 1.26 (79)	0.83 ± 0.43 (73)	1.25 ± 0.59 (66)	1.70 ± 0.42 (62)	1.29 ± 0.93 (56)
2	1.72 ± 1.36 (36)	2.22 ± 1.64 (36)	1.94 ± 1.22 (34)	1.92 ± 0.94 (34)	0.98 ± 0.45 (41)	1.42 ± 0.64 (41)	1.22 ± 0.43 (38)	1.16 ± 0.49 (33)

[a] *Number of CYP3A5*3, *6, and *7 alleles.* [b] *Number of copies of the ABCB1 C/G/C (wild-type) haplotype.*

Source: Data from Santoro et al. [61], expressed as mean ± SD (number of observations).

provided the drive for our studies of the impact of *CYP3A* and *ABCB1* polymorphisms on the dose-adjusted trough concentrations (C_0/dose) of CSA and TAC in blood during the first three months after renal transplantation [61,62].

A total of 301 clinically stable patients were enrolled. Basic immunosuppression consisted of a combination of CSA/azathioprine (AZA, n = 97), CSA/sirolimus (SRL, n = 53), TAC/AZA (n = 101), or TAC/SRL (n = 50). Patient demographic characteristics and a complete description of the study protocol are presented in the original manuscripts. Fasting blood samples were obtained from all patients immediately before the scheduled morning dose, the C_0 of CSA and TAC was measured, and dose-adjusted trough concentrations (C_0/dose; ng.ml^{-1}/mg) in the first week (±3 days) and at 1, 2, and 3 months (±5 days) following renal transplantation were used for the statistical analyses. The target C_0 for CSA was 250–300 ng/mL during the first week, 200–250 ng/mL during the first and second months, and 100–150 ng/mL during the third month. For TAC, the target C_0 concentrations were 12–15 ng/mL in the first week, 10–12 ng/mL during the first month, and 8–10 ng/mL during the second and third months.

In our first study [61], all patients were genotyped for the defective *CYP3A5*3* (rs776746), *CYP3A5*6* (rs10264272), and *CYP3A5*7* (rs41303343) polymorphisms and for the *ABCB1* 1236C>T (rs1128503), 2677G>T/A (rs2032582), and 3435C>T rs1045642) SNPs. *CYP3A5* and *ABCB1* haplotypes comprising these alleles were input using the *haplostats* software [63]. We explored the effects of *CYP3A5* and *ABCB1* haplotypes, demographical, and clinical characteristics on the response variables CSA and TAC C_0/dose by fitting longitudinal mixed effects models to the data. The final multivariate model included covariates that achieved a significance level of $P < 0.05$.

The daily doses of CSA and TAC required to reach the target C_0 concentrations varied, respectively, 5- to 7.5-fold and 5.7- to 16-fold among patients at each observation time (1 week, 1, 2, and 3 months). Table 45.5 shows the CSA and TAC C_0/dose at these time points, according to the number of *CYP3A5* defective alleles and the number of copies of the wild-type *ABCB1* haplotype (C/G/C) at the 1236C>T, 2677G>T/A, and 3435C>T loci. In the longitudinal univariate analyses, the number of copies of the wild-type *ABCB1* haplotype—but not the number of variant *CYP3A5* alleles—associated with CSA C_0/dose throughout the observation period. However, this association was not confirmed in the multivariate analysis.

By contrast, the longitudinal univariate analyses revealed a highly significant association ($P < 0.0001$) between the TAC C_0/dose and the number of variant *CYP3A5* alleles; none of the other variables examined, whether pharmacogenetic, clinical, or demographic, were found to be associated with the TAC C_0/dose. These results precluded a multivariate analysis. The increasingly negative regression coefficients for the association between TAC C_0/dose and *CYP3A5*, as the number defective alleles decrease from two (default) to one (regression coefficient = −1.299) and then to 0 (regression coefficient = −1.907), pointed to a gene–dose effect consistent with the inferred CYP3A5 activity phenotype; that is, as the number of the defective

*CYP3A5*3*, **6*, and **7* alleles decrease, CYP3A5-mediated TAC clearance increases, requiring higher TAC doses to reach the target C_0 blood levels, which in turn leads to lower TAC C_0/dose values.

Although many studies have shown a clear association between the *CYP3A5*1/*3* polymorphism and TAC pharmacokinetics, our investigation disclosed two novel and important findings. First, we extended this PGx association to two other nonfunctional *CYP3A5* alleles, namely *CYP3A5*6* and *CYP3A5*7* which are common in black Africans and African-derived populations. A clinically relevant implication of these findings is that in populations of African descent, inference of CYP3A5 metabolic activity requires genotyping for *CYP3A5*6* and *CYP3A5*7* in addition to *CYP3A5*3*. Second, the inclusion of Brazilian patients with the *CYP3A5*1/*1* genotype, which is rare or absent in populations of European or Asian descent, allowed demonstration of a gene–dose effect of *CYP3A5* polymorphisms on TAC pharmacokinetics.

Our finding that the association of *CYP3A5* genotype and TAC C_0/dose was present throughout the observation period, including the first week after transplantation, supports the strategy, adopted in some transplant centers, of genotyping for *CYP3A5* polymorphisms prior to transplantation [64]. Of note, a recent prospective study showed that adaptation of the initial TAC-loading dose according to the patient´s *CYP3A5* genotype led to a faster achievement of TAC target blood concentration and increased the number of patients within the desired TAC target range early after transplantation [65].

We have recently extended our PGx studies of immunosuppressants to two *CYP3A4* polymorphisms, namely *CYP3A4*1B* (rs2740574) and *CYP3A4*22* (rs35599367), which were reported to associate with TAC pharmacokinetics and dose requirements in renal allograft recipients [58–60]. We genotyped the TAC patients in our Brazilian cohort for the *CYP3A4*1B* and *CYP3A4*22* SNPs and repeated the two-step (univariate followed by multivariate) regression modeling of the association between TAC C0/dose at three months after transplantation and *CYP3A* polymorphisms, including these two *CYP3A4* variants, in addition to the previously investigated C*YP3A5*3*, **6*, and **7* alleles. We observed strong linkage disequilibrium (LD) between *CYP3A4*1B* and *CYP3A5*3* ($D' = 0.798$) but no LD for any other allele pair.

Univariate linear regression analysis revealed no association between TAC C_0/dose and *CYP3A4*22* but highly significant associations with *CYP3A5* haplotype and *CYP3A4*1B* genotype. The latter two variables were entered into a bivariate regression model, which retained only *CYP3A5* haplotype as a significant predictor of TAC C_0/dose ($P < 0.0001$; see Santoro et al. [62]). The exclusion of *CYP3A4*1B* from the final model was ascribed to LD between *CYP3A41B* and *CYP3A5*3*.

In summary, we confirmed the significant impact of *CYP3A5* haplotypes—comprising the *CYP3A5*3*, *CYP3A5*6*, and *CYP3A5*7* alleles—on TAC C_0/dose in Brazilian renal allograft recipients but did not verify independent effects of *CYP3A41B* or *CYP3A4*22* on TAC dose requirements at three months after transplantation. Differences in the study cohorts and/or clinical settings may account for this discrepancy. This possibility supports the notion that extrapolation of PGx data across populations requires caution and may be inappropriate, pointing to the importance of performing clinical PGx trials in the target populations.

45.4 PGX OF ANTIPSYCHOTICS IN SCHIZOPHRENIC PATIENTS

Psychiatric disorders are common and have high levels of morbidity and mortality. They strike early in life and are usually lifelong and chronic disorders. Although there are a variety of pharmacological treatments available, response and tolerability to drugs are highly variable. In general, response and remission rates to pharmacological treatment are limited. There are several potential explanations for these poor drug response rates, including clinical heterogeneity and diagnostic uncertainty; comorbidity; and environmental, social, and genetic factors. PGx studies have been carried out with multiple classes of psychotropic drugs, with a primary focus on the interindividual variation in drug efficacy.

Most investigations performed in Brazil have used a candidate gene approach, usually based on the receptor pharmacology of the psychotropic agent. Antipsychotic drugs are the mainstay of schizophrenia treatment. When they were developed during the 1950s, typical neuroleptic drugs revolutionized the treatment of psychosis. These medications control the excitation, aggression, and restlessness as well as conceptual disorganization and thought disorders of psychosis in a considerable proportion of cases; however, they are limited in their use by their relatively narrow therapeutic index, especially because of extrapyramidal side effects. The Brazilian public health system considers the typical neuroleptics as the first choice to treat schizophrenia, due to their lower acquisition costs, availability of depot formulations and minimal metabolic risks, improving adherence to drug therapy and offering an affordable alternative.

Table 45.6 summarizes the major outcomes of pharmacogenomics studies on antipsychotics on Brazilians carried out by Mara H. Hutz and colleagues. Kolhrausch et al. [66] explored the influence of SNPs in *DRD2* and *DRD3* genes, as well as polymorphisms in the *CYP3A* subfamily (*CYP3A4* and *CYP3A5*), and the *CYP2D6* gene on typical neuroleptic response in a Brazilian sample of patients with schizophrenia. The schizophrenia diagnosis was based on

TABLE 45.6 Main Findings on Antipsychotics Response in Brazilian Patients

Drug	Gene	Outcome	p value	Reference
Typical neuroleptics		Treatment refractoriness		
	DRD2		NS	[66]
	DRD3		0.021	
	CYP3A4		0.014	
	CYP3A5		0.003	
	CYP2D6		NS	
Clozapine		Clinical response		
	GNB3		0.021	[70]
	5-HTT		0.03	[71]
		Tonic-clonic seizures		
	GNB3		0.018	[70]
	CYP1A2		0.033	[80]

Diagnostic and Statistical Manual of Mental Disorder-IV criteria. All patients were of European ancestry. The nonrefractory group (35%) consisted of patients who had experienced adequate and long-lasting response to treatment with typical neuroleptics, most patients receiving comedication (e.g., anticholinergics, lithium, or antidepressants).

The criteria for treatment resistance included treatment refractoriness (inadequate clinical response) and treatment intolerance (adverse effects). The *CYP3A4*–392A>G SNP was significantly associated with efficacy of neuroleptic treatment. Homozygous individuals for the –392A variant ($p=0.014$, $OR=3.32$) were more frequent in the treatment-resistant group compared with carriers of one copy of the –392G variant. The *CYP3A5* low-expressor genotype (*CYP3A5**3/*CYP3A5**3) was found to be associated with refractoriness to neuroleptic treatment ($P=0.003$, $OR=3.16$). Among the haplotypes observed in the *DRD3* gene, the T/A/G/A/C haplotype showed an association with refractoriness to neuroleptics ($P=0.021$, $OR=1.75$). This association showed that carriers of one copy of this haplotype presented intermediate values between noncarriers and homozygous individuals for the haplotype.

No association was observed with polymorphisms in the *DRD2* and *CYP2D6* genes. Multiple logistic regression analyses showed that the number of copies of *DRD3* T/A/G/A/C haplotype and *CYP3A5* low-expressor genotype were predictors of refractoriness to neuroleptics after controlling for selected risk factors. *CYP3A5*3* homozygous individuals carrying at least one copy of the *DRD3* T/A/G/A/C haplotype showed a higher risk to be refractory to neuroleptics than *CYP3A5*3* homozygotes *DRD3* non-T/A/G/A/C carriers ($P=0.019$, $OR=2.32$; 95% confidence interval = 1.08–5.02). These results suggest a role for *CYP3A5* and *DRD3* gene variants on refractoriness to neuroleptic treatment in Brazilians with schizophrenia; nevertheless, the effect sizes are small.

Among the repertoire of drugs used to treat the disorder, several meta-analyses have evidenced that the antipsychotic clozapine displays greater effectiveness compared to typical neuroleptic drugs. Although clozapine is associated to a superior clinical efficacy than other antipsychotics, only 30–60% of individuals resistant to typical antipsychotics respond to it. At the same time, clozapine displays undesirable adverse effects, such as agranulocytosis, weight gain, hypotension, somnolence, constipation, and several types of convulsive symptoms, including new-onset-generalized seizures [67–71]. Clozapine has a unique receptor-binding profile, which includes several target affinities, ranging from low to high, with dopamine D1-5; serotonin 5-HT1-3, 5-HT_6, and 5-HT_7; α_{1-2}, and β_{1-3}-adrenergic; muscarinic cholinergic M_{1-5}; and histaminergic H_1 and H_3 receptors [72].

Since genetic factors are believed to play a significant role in the variation of response to antipsychotics, and these drugs exert their therapeutic effects by competitive antagonism of postsynaptic GPCRs, Kolhrausch et al. [70] evaluated the effect of 825C>T polymorphism in the *GNB3* gene on the response of clozapine treatment in a Brazilian sample of patients of European ancestry with *DSM-IV* diagnosis of schizophrenia. An association between the *GNB3* 825C>T and clinical response to clozapine was observed. Homozygosis for the variant 825T allele was more frequent among nonresponders (27.3%) than among responders (10.6%), whereas homozygosis for wild-type 825C allele was observed with higher frequency among responders (47.0% vs. 27.3% among nonresponders) ($\chi^2=7.708$, $p=0.021$). When we pooled homozygous and heterozygous individuals for the 825T allele (825T carriers), significant difference between responders and nonresponders ($\chi^2=4.939$, $p=0,026$, $OR=2.36$ [95%CI 1.03–5.50]) was observed.

Since the serotonin transporter 5HTTLPR polymorphism is likely to influence 5HT concentrations at all synapses, allelic variation in 5HTT function may affect the response to almost any agent affecting the 5HT system, especially atypical antipsychotics that have a higher occupancy rate of serotonin receptors. The HTTLPR polymorphism has been found to be associated with antipsychotic drug response in several studies (reviewed in Wilffert, Zaal, and Brouwers [73]). Kolhrausch et al. [71] examined the possible association between polymorphisms in the *5-HTT* gene polymorphisms and response to clozapine in South Brazilian patients. Significant differences between responders and nonresponders to clozapine were observed in the 5′ regulatory region of the human *5-HTT*. The 5HTTLPR S allele was more frequent in patients who failed to respond to clozapine (60%) when compared to those who did respond (43%). Logistic regression analysis confirmed this association after controlling for confounding variables ($OR=3.15$; 95% $CI=1.13$–8.80; $p=0.03$).

New-onset-generalized seizures represent a great burden in schizophrenic patients' care and a substantial impact on medication adherence and response. Clozapine lowers the threshold for generalized seizures [74] but the exact mechanism is not fully understood. Kolhrausch et al. [69] observed that carriers of the *GNB3* 825T allele showed an increased risk for convulsive episodes ($\chi^2=5.607$, $p=0.018$); they also found that this association was independent of clozapine dose. The most plausible explanation for these findings is that G protein-activated inwardly rectifying K^+ (GIRK) channels play an important role in the inhibitory regulation of neural excitability in most brain regions through activation of various GPCRs [75].

It has been shown that GIRK channels are involved in seizure susceptibility [76], and dose-related inhibition of this function by antipsychotics can increase seizure susceptibility. Since clozapine strongly inhibits GIRK channels [77], this effect can explain the relationship between *GNB3* polymorphisms and seizures in patients exposed to this drug. Two recent studies [78,79] suggested that specific combinations of *CYP1A2* *1F and *1C alleles could induce lower messenger RNA (mRNA) expression and predispose to clozapine intolerance. In an attempt to partially corroborate this finding, we evaluated these two polymorphisms in a sample of South Brazilians with schizophrenia and their importance in the proconvulsive effect of clozapine [80]. We found the *1F/*1F genotype to be significantly associated with seizures under clozapine treatment ($p=0.033$, $OR=2.69$, 95% $CI=1.042$–6.940). Logistic regression analysis confirmed the importance of this genotype. No relationship was observed between combinations of *1F and *1C alleles.

At present, the challenge faced by research into the genetic basis of psychopharmacological drug responses is to identify genes with a relative small effect against a background of substantial genetic and environmental variations. Additionally, selection bias and factors, such as socioeconomic status or healthcare, may differ between individuals. Differences in studies' results may be due to heterogeneity among experimental protocols (i.e., treatment protocols, length of follow up, patient selection, allele classifications), as well as differences in allele frequencies between populations.

45.5 CONCLUSIONS

The kaleidoscopic diversity of the Brazilian population, with trihybrid biogeographical ancestry in Europe, Africa, and America affects the distribution of PGx polymorphisms and poses a caveat to extrapolation of PGx data from comparatively more homogeneous populations. These conditions prompted the creation of the Brazilian Pharmacogenetics Network, Refargen, the mission of which is to provide leadership in PGx research and education in Brazil and make an impact on population health.

A major challenge to PGx studies in Brazil is population stratification, which if not controlled for, may confound the outcomes of PGx association studies. Refargen researchers have developed ways to control for this challenge by combining ancestry-informative markers and appropriate statistical approaches. A distinct message that emerges from these analyses is that race/Color categorization does not capture the distribution of PGx polymorphisms among Brazilians, which is best modeled by continuous functions of the individual proportions of European and African ancestry, irrespective of self-identified Color. Recognition of this fact is important in the design and interpretation of PGx clinical trials in Brazilians but does not imply that PGx-informed drug prescription requires investigation of individual ancestry. Rather, individual genotyping should be directed to PGx polymorphisms of proven clinical utility for the specific medical condition being treated, irrespective of biogeographical ancestry.

Refargen groups have conducted PGx clinical trials involving a range of therapeutic medicines, and in the case of warfarin, derived dosing algorithms for clinical implementation. Adoption of these algorithms in the natural clinical setting of the Brazilian public health system and private practice faces the same barriers encountered worldwide, not only for warfarin but for PGx-informed prescription in general. The Refargen consortium contributes to overcoming these barriers through research, education, training, and consulting initiatives; these will benefit not only Brazilians but also other admixed Latin American peoples, who are infrequently enrolled in drug development programs by the international pharma industry.

ACKNOWLEDGMENTS

G. S-K is supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (Faperj), and Financiadora de Estudos e Projetos (Finep). M.H.H is funded by CNPq and Finep.

REFERENCES

[1] Suarez-Kurtz G, Pena SDJ. Pharmacogenomics in the Americas: impact of genetic admixture. Curr Drug Targets 2006;7:1649–58.

[2] Pena SDJ, Di-Pietro G, Fuchshuber-Moraes M, Pasqualini-Genro J, Hutz MH, Kehdi F, et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. PLoS One 2011;6(2):e17063.

[3] Suarez-Kurtz G. Pharmacogenetics in the Brazilian population. Front Pharmacol 2010;1:118.

[4] Suarez-Kurtz G, Pena SD, Struchiner CJ, Hutz MH. Pharmacogenomic diversity among Brazilians: influence of ancestry, self-reported color, and geographical origin. Front Pharmacol 2012;3:191.

[5] Suarez-Kurtz G, Pena SDJ. Pharmacogenetic studies in the Brazilian population. In: Suarez-Kurtz G, editor. Pharmacogenomics in admixed populations. Landes Bioscience; 2007. p. 75–98.

[6] Suarez-Kurtz G. Pharmacogenomics in admixed populations, the Brazilian pharmacogenetics/pharmacogenomics network–REFARGEN. Pharmacogenomics J 2004;4:47–8.

[7] Rede Nacional de Farmacogenética, Refargen. Rio de Janeiro; 2013. <www.refargen.org.br> [homepage on the Internet].

[8] Suarez-Kurtz G. Conference scene: Latin American pharmacogenomics and personalized medicine conference. Pharmacogenomics 2012;13:1449–52.

[9] Suarez-Kurtz G, Amorim A, Damasceno A, Hutz MH, Moraes MO, Ojopi EB, et al. *VKORC1* polymorphisms in Brazilians, comparison with the Portuguese and Portuguese-speaking Africans and pharmacogenetic implications. Pharmacogenomics 2010;11:1257–67.

[10] Suarez-Kurtz G, Genro JP, Moraes MO, Ojopi EB, Pena SDJ, Perini JA, et al. Global pharmacogenomics: impact of population diversity on the distribution of polymorphisms in the CYP2C cluster among Brazilians. Pharmacogenomics J 2012;12:67–76.

[11] Sortica Vde A, Ojopi EB, Genro JP, Callegari-Jacques S, Ribeiro-Dos-Santos A, de Moraes MO, et al. Influence of genomic ancestry on the distribution of SLCO1B1, SLCO1B3 and ABCB1 gene polymorphisms among Brazilians. Basic Clin Pharmacol Toxicol 2012;110:460–8.

[12] Rede Nacional de Farmacogenética, Refargen. Rio de Janeiro; 2013. <www.refargen.org.br/rubrique.php3?id_rubrique=28> [homepage on the Internet].

[13] Salzano FM, Bortolini MC. The evolution and genetics of Latin American populations. Cambridge University Press; 2002.

[14] Bastos-Rodrigues L, Pimenta JR, Pena SDJ. The genetic structure of human populations studied through short insertion-deletion polymorphisms. Ann Hum Genet 2006;70:658–65.

[15] Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, et al. Color and genomic ancestry in Brazilians. Proc Natl Acad Sci USA 2003;100:177–82.

[16] Suarez-Kurtz G, Perini JA, Bastos-Rodrigues L, Pena SDJ, Struchiner CJ. Impact of population admixture on the distribution of the CYP3A5*3 polymorphism. Pharmacogenomics 2007;8:1299–306.

[17] Estrela RCE, Ribeiro FS, Carvalho RS, Gregorio SP, Dias-Netto E, Struchiner CJ, et al. The distribution of *ABCB1* polymorphisms among Brazilians: impact of population admixture. Pharmacogenomics 2008;9:267–76.

[18] Instituto Brasileiro de Geografia e Estatística, IBGE. Brasilia; 2013. <www.sidra.ibge.gov.br/bda/tabela/protabl.asp?c=3145&z=cd&o=4&i=P> [homepage on the Internet].

[19] Suarez-Kurtz G, Pena SDJ, Hutz MH. Application of the F(ST) statistics to explore pharmacogenomic diversity in the Brazilian population. Pharmacogenomics 2012;13:771–7.

[20] Wright S. Genetical structure of populations. Nature 1950;166:247–9.

[21] Wright S. Variability within and among natural populations. University of Chicago; 1978.

[22] Suarez-Kurtz G, Vargens DD, Struchiner CJ, Bastos-Rodrigues L, Pena SDJ. Self-reported skin color, genomic ancestry and the distribution of GST polymorphisms. Pharmacogenet Genomics 2007;17:765–71.

[23] Vargens DD, Almendra L, Struchiner CJ, Suarez-Kurtz G. Distribution of the GNB3 825C>T polymorphism among Brazilians, impact of population structure. Eur J Clin Pharmacol 2008;3:253–6.

[24] Rodrigues AC, Hirata MH, Hirata RD. The genetic determinants of atorvastatin response. Curr Opin Mol Ther 2007;9:545–53.

[25] Rodrigues AC, Perin PM, Purim SG, Silbiger VN, Genvigir FD, Willrich MA, et al. Pharmacogenetics of OATP transporters reveals that SLCO1B1 c.388A>G variant is determinant of increased atorvastatin response. Int J Mol Sci 2011;12:5815–27.

[26] Rodrigues AC, Sobrino B, Genvigir FD, Willrich MA, Arazi SS, Dorea EL, et al. Genetic variants in genes related to lipid metabolism and atherosclerosis, dyslipidemia and atorvastatin response. Clin Chim Acta 2013;417:8–11.

[27] Luchessi AD, Silbiger VN, Cerda A, Hirata RD, Carracedo A, Brion M, et al. Increased clopidogrel response is associated with ABCC3 expression: a pilot study. Clin Chim Acta 2012;413:417–21.

[28] Piranda DN, Festa-Vasconcellos JS, Amaral LM, Bergmann A, Vianna-Jorge R. Polymorphisms in regulatory regions of cyclooxygenase-2 gene and breast cancer risk in Brazilians: a case-control study. BMC Cancer 2010;10:613.

[29] Festa-Vasconcellos JS, Piranda DN, Amaral LM, Indio-do-Brasil V, Koifman S, Vianna-Jorge R. Polymorphisms in cycloxygenase-2 gene and breast cancer prognosis: association between PTGS2 haplotypes and histopathological features. Breast Cancer Res Treat 2012;132:251–8.

[30] Vianna-Jorge R, Festa-Vasconcellos JS, Goulart-Citrangulo SM, Leite MS. Functional polymorphisms in xenobiotic metabolizing enzymes and their impact on the therapy of breast cancer. Front Genet 2012;3:329.

[31] Turatti L, Sprinz E, Lazzaretti RK, Kuhmmer R, Agnes G, Silveira JM, et al. Short communication: UGT1A1* 28 variant allele is a predictor of severe hyperbilirubinemia in HIV-infected patients on HAART in southern Brazil. AIDS Res Hum Retroviruses 2012;28:1015–8.

[32] Possuelo LG, Castelan JA, de Brito TC, Ribeiro AW, Cafrune PI, Picon PD, et al. Association of slow N-acetyltransferase 2 profile and anti-TB drug-induced hepatotoxicity in patients from Southern Brazil. Eur J Clin Pharmacol 2008;64:673–81.

[33] Santos NP, Callegari-Jacques SM, Ribeiro Dos Santos AK, Silva CA, Vallinoto AC, Fernandes DC, et al. N-acetyl transferase 2 and cytochrome P450 2E1 genes and isoniazid-induced hepatotoxicity in Brazilian patients. Int J Tuberc Lung Dis 2013;17:499–504.

[34] Vianna-Jorge R, Perini JA, Rondinelli E, Suarez-Kurtz G. *CYP2C9* genotypes and the pharmacokinetics of tenoxicam in Brazilians. Clin Pharmacol Ther 2004;76:18–26.

[35] Perini JA, Vianna-Jorge R, Brogliato AR, Suarez-Kurtz G. Influence of *CYP2C9* genotypes on the pharmacokinetics and pharmacodynamics of piroxicam. Clin Pharmacol Ther 2005;78:362–9.

[36] Perini JA, Suarez-Kurtz G. Impact of CYP2C9*3/*3 genotype on the pharmacokinetics and pharmacodynamics of piroxicam. Clin Pharmacol Ther 2006;80:549–51.

[37] Ministério da Saúde, Departamento de DST, Aids e Hepatites virais. Brasília; 2013. <www.aids.gov.br> [homepage on the Internet].

[38] Estrela RCE, Ribeiro FS, Barroso PF, Tuyama M, Gregório SP, Dias-Neto E, et al. ABCB1 polymorphisms have no impact on the concentrations of lopinavir and ritonavir in blood, semen and saliva of HIV-infected men under stable antiretroviral therapy. Pharmacogenomics 2009;10:311–8.

[39] Estrela RCE, Santoro AB, Barroso PF, Tuyama M, Suarez-Kurtz G. *CYP3A5* genotype has no impact on plasma trough concentrations of lopinavir and ritonavir in HIV-infected subjects. Clin Pharmacol Ther 2008;84:205–7.

[40] Josephson F, Allqvist A, Janabi M, Sayi J, Aklillu E, Jande M, et al. CYP3A5 genotype has an impact on the metabolism of the HIV protease inhibitor saquinavir. Clin Pharmacol Ther 2007;81:708–12.

[41] Hartkoorn RC, Kwan WS, Shallcross V, Chaikan A, Liptrott N, Egan D, et al. HIV protease inhibitors are substrates for OATP1A2, OATP1B1 and OATP1B3 and lopinavir plasma concentrations are influenced by SLCO1B1 polymorphisms. Pharmacogenet Genomics 2010;20:112–20.

[42] Kohlrausch FB, Estrela RCE, Barroso PF, Suarez-Kurtz G. The impact of *SLCO1B1* polymorphisms on the plasma concentration of lopinavir and ritonavir in HIV-infected men. Br J Clin Pharmacol 2010;69:95–8.

[43] Rede Nacional de Farmacogenética, Refargen. Rio de Janeiro; 2013. <www.refargen.org.br/article.php3?id_article=55> [homepage on the Internet].

[44] Limdi NA, Wadelius M, Cavallari L, Eriksson N, Crawford DC, Lee MT, et al. Warfarin pharmacogenetics: a single VKORC1 polymorphism is predictive of dose across 3 racial groups. International warfarin pharmacogenetics consortium. Blood 2010;115:3827–34.

[45] Perini JA, Struchiner CJ, Silva-Assunção E, Santana ISC, Rangel F, Ojopi EB, et al. Pharmacogenetics of warfarin: development of a dosing algorithm for Brazilian patients. Clin Pharmacol Ther 2008;84:722–8.

[46] Perini JA, Struchiner CJ, Silva-Assunção E, Suarez-Kurtz G. Impact of CYP4F2 rs2108622 on the stable warfarin dose in an admixed patient cohort. Clin Pharmacol Ther 2010;87:417–20.

[47] Suarez-Kurtz G, Perini JA, Silva-Assunção E, Struchiner CJ. Relative contribution of VKORC1, CYP2C9 and INR response to warfarin stable dose. Blood 2009;113:4125–6.

[48] Botton MR, Bandinelli E, Rohde LE, Amon LC, Hutz MH. Influence of genetic, biological and pharmacological factors on warfarin dose in a Southern Brazilian population of European ancestry. Br J Clin Pharmacol 2011;72:442–50.

[49] Suarez-Kurtz G. Population diversity and the performance of warfarin dosing algorithms. Br J Clin Pharmacol 2011;72:451–3.

[50] Gage BF, Eby C, Johnson JA, Deych E, Rieder MJ, Ridker PM, et al. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. Clin Pharmacol Ther 2008;84:326–31.

[51] Limdi NA, Beasley TM, Crowley MR, Goldstein JA, Rieder MJ, Flockhart DA, et al. VKORC1 polymorphisms, haplotypes and haplotype groups on warfarin dose among African-Americans and European-Americans. Pharmacogenomics 2008;9:1445–58.

[52] Fuchshuber-Moraes M, Perini JA, Rosskopf D, Suarez-Kurtz G. Exploring warfarin pharmacogenomics with the extreme-discordant-phenotype methodology: impact of FVII polymorphisms on stable anticoagulation with warfarin. Eur J Clin Pharmacol 2009;65:789–93.

[53] International Warfarin Pharmacogenetics ConsortiumKlein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med 2009;360:753–64.

[54] Beutler E, Gelbart T, Demina A. Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? Proc Natl Acad Sci USA 1988;95:8170–4.

[55] Vargens DD, Neves RR, Bulzico DA, Ojopi EB, Meirelles RM, Pessoa CN, et al. Association of the UGT1A1-53(TA)n polymorphism with L-thyroxine doses required for thyrotropin suppression in patients with differentiated thyroid cancer. Pharmacogenet Genomics 2011;21:341–3.

[56] Haufroid V, Wallemacq P, VanKerckhove V, et al. CYP3A5 and ABCB1 polymorphisms and tacrolimus pharmacokinetics in renal transplant candidates: guidelines from an experimental study. Am J Transplant 2006;6:2706–13.

[57] Staatz CE, Goodman LK, Tett SE. Effect of CYP3A and ABCB1 single nucleotide polymorphisms on the pharmacokinetics and pharmacodynamics of calcineurin inhibitors: Part I. Clin Pharmacokinet 2010;49:141–75.

[58] Tavira B, Coto E, Díaz-Corte C, Ortega F, Arias M, Torres A, et al. Pharmacogenetics of tacrolimus after renal transplantation: analysis of polymorphisms in genes encoding 16 drug metabolizing enzymes. Clin Chem Lab Med 2011;49:825–33.

[59] Elens L, Bouamar R, Hesselink DA, Haufroid V, van der Heiden IP, van Gelder T, et al. A new functional CYP3A4 intron 6 polymorphism significantly affects tacrolimus pharmacokinetics in kidney transplant recipients. Clin Chem 2011;57:1574–83.

[60] de Jonge H, de Loor H, Verbeke K, Vanrenterghem Y, Kuypers DR. *In vivo CYP3A4* activity, *CYP3A5* genotype, and hematocrit predict tacrolimus dose requirements and clearance in renal transplant patients. Clin Pharmacol Ther 2012;92:366–75.

[61] Santoro A, Felipe CR, Tedesco-Silva H, Medina-Pestana JO, Struchiner CJ, Ojopi EB, et al. Pharmacogenetics of calcineurin inhibitors in Brazilian renal transplant patients. Pharmacogenomics 2011;12:1293–303.

[62] Santoro AB, Struchiner CJ, Felipe CR, Tedesco-Silva H, Medina-Pestana JO, Suarez-Kurtz G. CYP3A5 genotype, but not CYP3A4*1b, CYP3A4*22, or hematocrit, predicts tacrolimus dose requirements in Brazilian renal transplant patients. Clin Pharmacol Ther 2013. http://dx.doi.org/10.1038/clpt.2013.68. [Epub ahead of print.]

[63] Mayo Clinic, Statistical Genetics and Genetic Epidemiology Laboratory. Rochester; 2013. <http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm> [homepage on the Internet].

[64] MacPhee IA, Holt DW. A pharmacogenetic strategy for immunosuppression based on the CYP3A5 genotype. Transplantation 2008;85:163–5.

[65] Thervet E, Loriot MA, Barbier S, Buchler M, Ficheux M, Choukroun G, et al. Optimization of initial tacrolimus dose using pharmacogenetic testing. Clin Pharmacol Ther 2010;87:721–6.

[66] Kohlrausch FB, Gama CS, Lobato MI, Belmonte-De-Abreu PS, Callegari-Jacques SM, Gesteira A, et al. Naturalistic pharmacogenetic study of treatment resistance to typical neuroleptics in European-Brazilian schizophrenics. Pharmacogenet Genomics 2008;18:599–609.

[67] Davis JM, Chen N, Glick ID. A meta-analysis of the efficacy of second-generation antipsychotics. Arch Gen Psychiatry 2003;60:553–64.

[68] Arranz MJ, Munro J, Osborne S, Collier D, Kerwin RW. Applications of pharmacogenetics in psychiatry: personalisation of treatment. Expert Opin Pharmacother 2001;2:537–42.

[69] Basile VS, Masellis M, Potkin SG, Kennedy JL. Pharmacogenomics in schizophrenia: the quest for individualized therapy. Hum Mol Genet 2002;11:2517–30.

[70] Kohlrausch FB, Salatino-Oliveira A, Gama CS, Lobato MI, Belmonte-de-Abreu P, Hutz MH. G-protein gene 825C>T polymorphism is associated with response to clozapine in Brazilian schizophrenics. Pharmacogenomics 2008;9:1429–36.

[71] Kohlrausch FB, Salatino-Oliveira A, Gama CS, Lobato MI, Belmonte-de-Abreu PS, Hutz MH. Influence of serotonin transporter gene polymorphisms on clozapine response in Brazilian schizophrenics. J Psychiatr Res 2010;44:1158–62.

[72] Wilffert B, Zaal R, Brouwers JR. Pharmacogenetics as a tool in the therapy of schizophrenia. Pharm World Sci 2005;27:20–30.

[73] Vázquez-Bourgon J, Arranz MJ, Mata I, Pelayo-Terán JM, Pérez-Iglesias R, Medina-González L, et al. Serotonin transporter polymorphisms and early response to antipsychotic treatment in first episode of psychosis. Psychiatry Res 2010;175:189–94.

[74] Hedges D, Jeppson K, Whitehead P. Antipsychotic medication and seizures: a review. Drugs Today 2003;39:551–7.

[75] Blednov YA, Stoffel M, Alva H, Harris RA. A pervasive mechanism for analgesia: activation of GIRK2 channels. Proc Natl Acad Sci USA 2003;100:277–82.

[76] Signorini S, Liao YJ, Duncan SA, Jan LY, Stoffel M. Normal cerebellar development but susceptibility to seizures in mice lacking G protein-coupled, inwardly rectifying K+ channel GIRK2. Proc Natl Acad Sci USA 1997;94:923–7.

[77] Kobayashi T, Ikeda K, Kumanishi T. Effects of clozapine on the delta- and kappa-opioid receptors and the G-protein-activated K+ (GIRK) channel expressed in Xenopus oocytes. Br J Pharmacol 1998;123:421–6.

[78] Bolla E, Bortolaso P, Ferrari M, Poloni N, Callegari C, Marino F, et al. Are CYP1A2*1F and *1C associated with clozapine tolerability? A preliminary investigation. Psychiatry Res 2011;189:483.

[79] Ferrari M, Bolla E, Bortolaso P, Callegari C, Poloni N, Lecchini S, et al. Association between CYP1A2 polymorphisms and clozapine-induced adverse reactions in patients with schizophrenia. Psychiatry Res 2012;200:1014–7.

[80] Kohlrausch FB, Severino-Gama C, Lobato MI, Belmonte-de-Abreu P, Carracedo A, Hutz MH. The CYP1A2 -163C>A polymorphism is associated with clozapine-induced generalized tonic-clonic seizures in Brazilian schizophrenia patients. Psychiatry Res 2013;209:242–5.

[81] International HapMap project. Bethesda; 2013. <http://hapmap.ncbi.nlm.nih.gov/> [homepage on the Internet].

This page is intentionally left blank

Chapter 46

Pharmacogenomics in India

Umamaheswaran Gurusamy and Deepak Gopal Shewade
Pharmacogenomics Laboratory, Department of Pharmacology, Jawaharlal Institute of Postgraduate Medical Education and Research, India

Chapter Outline

46.1 Overview **1037**
46.2 Indian Population Structure and Diversity **1038**
46.3 Frequency Distribution and Pattern of Genetic Variants of Pharmacogenes **1039**
46.3.1 Drug-Metabolizing Genes 1039
46.3.2 Drug-Transporter Genes 1050
46.3.3 Drug-Target Genes 1051
46.4 Pharmacogenetic Studies in India **1052**
46.4.1 Anticancer Drugs 1052
46.4.2 Cardiovascular Drugs 1053
46.4.3 Oral Hypoglycemic Drugs 1054
46.4.4 Anticonvulsants 1055
46.4.5 Immunosuppressants 1055
46.4.6 Analgesics 1056
46.4.7 Antifolates 1056
46.5 Conclusions and Future Directions **1056**
References **1057**

46.1 OVERVIEW

In a given homogenous population, a significant proportion of patients confer resistance, some respond differently and some develop serious adverse effects when they are prescribed under the model of "one dose fits for all" medication [1]. The completion of the Human Genome Project and successive invention of comprehensive high throughput "omics" technologies has initiated a revolution in genomic medicine in the last few years [2]. Intriguingly, in spite of these advances in medical care, variability in drug efficacy and susceptibility to drug toxicity remains a major challenge in healthcare systems. Genetics not only plays a significant role in the elucidation of disease causation but also a major contributor to the disproportions in pharmacotherapy. Inherited (germline) or acquired (somatic) changes in the gene sequence that codes for the proteins involved in the absorption, distribution, metabolism, and excretion (ADME) of drugs are the principal cause of variability in treatment outcomes [3]. These genetic variations are now being applied to predict the dosage requirement, therapeutic response, and drug-induced toxicity in clinical medicine under the discipline pharmacogenetics and pharmacogenomics (PGx).

The goal of PGx-guided drug therapy lies in its ability to provide predictive, preventive, and personalized medicine with the application of genomic data in clinical practice. This aids researchers in understanding the causative genetic state for the altered pharmacokinetic and pharmacodynamic effect of a drug through assessment of an individual's genetic profile in conjunction with clinical and environmental covariates. Technological advances in human genetics have facilitated significant progress in PGx and stimulated its research from the level of single-nucleotide polymorphism (SNP) (monogenic) measurement to the comprehensive genome-wide approach (oligogenic). The adoption of new molecular-based methodologies and testing strategies has allowed both healthcare providers and pharmaceutical companies to expand and fine-tune the scope of clinical PGx beyond drug discovery, development, and clinical trials.

Interestingly, the concept and significance of interindividual variability was emphasized back in 460–370 BC by the Greek physician Hippocrates through his observation that "it is more important to know what sort of person has a disease than to know what sort of disease a person has." The hope of personalized medicine was foreseen in the 1800s by the Canadian physician Sir William Osler, who said, "variability is the law of life, and as no two faces are the same, so

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00046-3
© 2014 Elsevier Inc. All rights reserved.

no two bodies are alike, and no two individuals react alike and behave alike under the abnormal conditions we know as disease" [4]. Later, during the 1930s, the documentation of the racial variability in taste blindness for phenylthiocarbamide by Snyder marked the beginning of PGx [5]. Nevertheless, the term "pharmacogenetics" was first coined and published by the German physician Friedrich Vogel in 1959 [6]. Since then PGx has evolved into a major driving force as a result of a series of discoveries that have heightened its development.

As with most drugs, research focuses on identifying genetic markers that will predict the safety and effectiveness of a drug. Because of emerging genomic knowledge, the US Food and Drug Administration (FDA) now requires PGx-related information and warnings on drug labels [7]. The scope of this present chapter is intended to provide up-to-date information on the status of pharmacogenomics research in the Indian population. We focus on the pharmacogenetic studies in our laboratory on drugs such as anticancer drugs, antiestrogen agents, antiplatelets, anticoagulants, HMG-COA reductase inhibitors, oral hypoglycemic medications, anticonvulsants, and analgesics. Further, we summarize the data from other therapeutic groups in India and discuss the challenges, current controversies, knowledge gaps, and strategies surrounding the application of PGx testing in personalized medicine for Indians.

46.2 INDIAN POPULATION STRUCTURE AND DIVERSITY

India is the seventh largest country by area, occupying around 3,287,263 km^2 and contributing about one-sixth of the global population. Presently it is home to more than 1.2 billion and is the most peopled nation in the world after China [8]. The Indian population is encompassed of numerous subgroups that are extensively diversified by their genetic constitution, language, geography, and religious, sociocultural, and ancestry components. These distinguished features together render them a unique resource of PGx heterogeneity compared to the rest of the world. Today, there are 6 religions (Hinduism, Islam, Christianity, Sikhism, Buddhism, and Jainism), 4693 different communities with many thousands of endogamous groups, 325 functioning languages, and 25 scripts in the Indian subcontinent [9].

Modern-day Indians are morphologically classified into four primary ethnic strains and belong to the following groups: Caucasoid, Mongoloid, Proto-Australoid, and Negrito (Figure 46.1). The Caucasoids and Proto-Australoids predominate and are mostly confined to Northern and Southern India, respectively. The Mongoloids live along the Himalayan fringe of Jammu and Kashmir and in the northeastern region of the country, while the Negritos are restricted to the Andaman and Nicobar Islands.

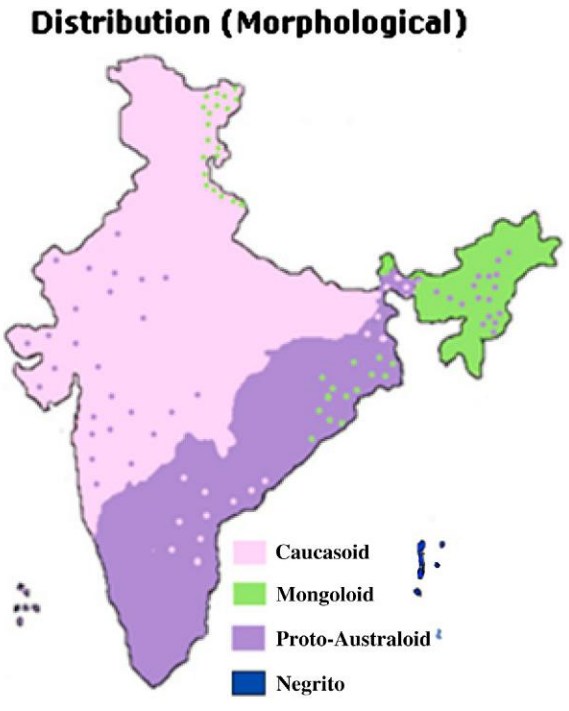


FIGURE 46.1 Morphological distribution of Indian populations. (*Reprinted from "The Indian Genome Variation database (IGVdb): a project review by Indian Genome Variation Consortium, 2005 Hum Genet"*)

The indigenous Negritos are believed to be the descendents of the first humans who migrated out of Africa about 65,000 years ago. The languages spoken by Indians have been broadly classified into four major linguistic families, as shown in Figure 46.2: Austro-Asiatic (Nishada) speakers are tribals and believed to be the most ancient inhabitants; their dialect is spoken by people mostly in Central, Eastern, and Northeastern India. Dravidian (Dravida) is predominantly spoken by Southern Indians. Indo-European (Aryan) is widely spoken in Northern India. And Tibeto-Burman (Kirata) is limited to the population of Northeast India, which is made up of immigrants from Tibet and Myanmar.

As a consequence of different waves of migration and invasion at different periods in history, India is also peopled by several recently drifted populations, and this has led to the admixture of various racial groups within the innate populations. Perhaps nowhere else in the world has the magnitude of genetic diversity in India been observed except for Africa. The contemporary populations are the intermixture of two contrasting ancestral groups—the ancestral North Indians (ANI) and the ancestral South Indians (ASI). About 30–70% of the ANI are genetically identical to the populations of the Middle East, Central Asia, and Europe. Conversely, the South Indians are Dravidian speakers who do not share their genetic signatures with any other global racial groups [10].

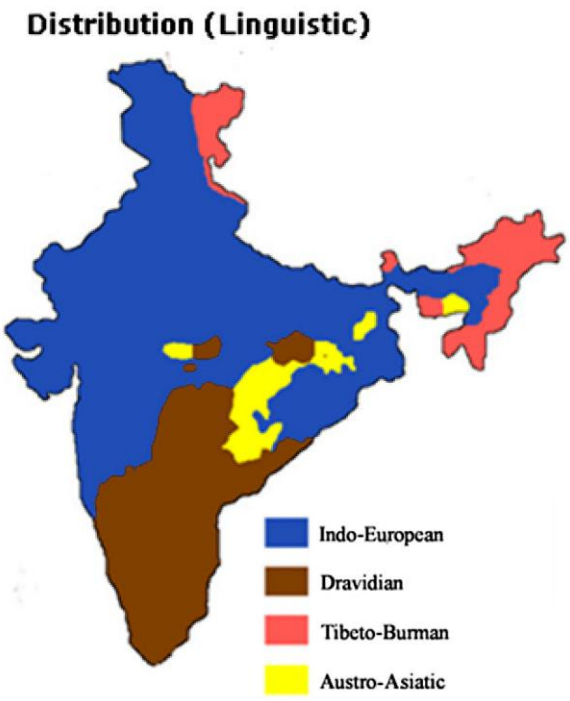


FIGURE 46.2 Linguistic distribution of Indian populations. (*Reprinted from "The Indian Genome Variation database (IGVdb): a project review by Indian Genome Variation Consortium, 2005 Hum Genet"*)

46.3 FREQUENCY DISTRIBUTION AND PATTERN OF GENETIC VARIANTS OF PHARMACOGENES

The frequency of polymorphic variants in ADME pharmacogenes are well known and exhibit wide inter- and intraethnic variability among global populations. Existing evidence suggests that knowledge of genetic variations is essential to better understand ethnic differences in disease occurrence, development, prognosis, therapeutic response, and toxicity. In addition, population-specific information regarding the normative frequency distribution of pharmacogenetic variants can serve as a basis for defining sample sizes for genotype–phenotype association studies.

Current studies in India have established the normative frequency of clinically important pharmacogenes, but most of them have analyzed a limited number of samples from South and North Indian populations [11–14]. Keeping in mind the extent of genetic heterogeneity, the results from these studies cannot be extrapolated to the entire Indian population. Recently, we compiled normative frequency data on genes coding proteins involved in drug metabolism and drug transport in various geographical regions of India and reported from different studies. These data have been pooled as North Indians (NI), South Indians (SI), and North East Indians (NEI) based on these population's ancestral ethnicity [11]. In this comprehensive review, we found that there is racial delineation between Indians and other major global populations. Also, we found significant intraethnic differences in the distribution of the molecular variants of ADME genes. In a previous study [15], we reported significant heterogeneity in the frequencies of *UGT1A1* and *MDR1* alleles among the four South Indian subpopulations.

Interestingly, allele frequency differentiation among Indian populations is greater than that of European populations. This is due to the strong founder effect and strict endogamy practice [10]. Table 46.1 presents normative frequency data on pharmacogenetic variants in ADME pharmacogenes that modulate drug action at various levels in Indian populations. These variants include *CYP1A1*, *CYP1A2, CYP2A6*, *CYP2C8*, *CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP2E1*, *CYP3A4*, *CYP3A5*, *GSTT1*, *GSTM1*, *GSTP1*, *UGT1A1*, *UGT1A7*, *TPMT*, *SULT1A1*, *COMT*, *NAT2*, *MDR1*, *OCT1*, *SLCO1B1*, *CYP19A1*, *VKORC1*, *MTHFR*, and *ACE*.

46.3.1 Drug-Metabolizing Genes

46.3.1.1 CYP1A1

CYP1A1 is an important member of the CYP1A family that is responsible for the metabolic activation of procarcinogens (polycyclic aromatic hydrocarbons and aromatic amines) into reactive metabolites. It also participates in the metabolism of steroidal hormones including estrogens. The *CYP1A1* gene is mapped on chromosome 15q22-q24, spanning 5810 bp and composed of 7 exons and 6 introns. Though several polymorphisms have been previously identified for the *CYP1A1* gene, much attention has been given to *CYP1A1 *2A* or *m1* (also known as *Msp1*), *CYP1A1*2C* or *m2*, *CYP1A1*3* or *m3* and *CYP1A1*4* or *m4* polymorphisms.

Genetic polymorphisms of the *CYP1A1* gene alters the enzyme activity and has been shown to be associated with different cancers such as carcinoma of breast, colon, ovary, lung, oral cavity and acute lymphoblastic leukemia (ALL). The prevalence of polymorphic alleles *CYP1A1*2A* and *CYP1A1*2C* is significantly different among the Indian populations and also in comparison with other populations. On the other hand *CYP1A1*3* is absent in Indians, while *CYP1A1*2B* (9.5%) is seen only in NEI, which is significantly different from other populations. The frequency of *CYP1A1*4* in NEI (40.5%) is different from that in NI (3%), Caucasians (4.3%) and Africans (1.2%), but this allele is absent in Asians. Conversely, *CYP1A1*4* frequency in NI matches that in Caucasians but is different from that in Africans [11].

46.3.1.2 CYP1A2

Cytochrome P4501A2 (CYP1A2) is involved in the bioactivation of procarcinogens into carcinogenic substances. In addition, CYP1A2 contribute to the metabolism of

TABLE 46.1 Normative Allele Frequency Distribution of Genes Encoding Drug Metabolism, Transporter, and Target Proteins in Indian Populations Compared to that in Other Major Populations

Gene and SNP	Location	Effect	dbSNP ID	Minor Allele Frequency					
				Africans	*Asians*	*Caucasians*	*NI*[*]	*SI*[*]	*NEI*[*]
CYP1A1									
*m1 or *2A (3798T>C)*	3′ UTR	Increased	rs4646903	21.8[a,c]	18.9[a,c]	8.5[a,b,c]	30.3[b,c]	21.6[a,c]	33.8[a,b]
**2B (3798T>C) (2455A>G)*	3′ UTR, Exon 7	Ile462Val	–	1.8[c]	21.2[c]	1.3[c]	NA	na	9.5
*m2 or *2C (2455A>G)*	Exon 7	Ile462Val	rs1048943	0.7[a,b,c]	32.8[a,b,c]	0.2[a,b,c]	22.0[b,c]	16.0[a,c]	11.5[a,c]
*m3 or *3 (3204T>C)*	3′ UTR	–	rs4986884	9.3	0	0.4	0	na	na
*m4 or *4 (2452C>A)*	Exon 7	Thr461Asn	rs1799814	1.2[a,c]	0	4.3[c]	3.0[c]	na	40.5[a]
CYP1A2									
**1C (−3860G>A)*	5′ UTR	Decreased	rs2060514	0.8[a,+]	21.1[a]	51.7[a,+]	8.0	na	na
**1D (2467Tdel)*	5′ UTR	–	rs35694136	na	42.0[a]	30.2	24.5	na	na
**1E (−729T>G)*	Intron 1	–	rs2069526	12.8[+]	8.2	0.4[a,+]	9.5	na	na
**1F (−163C>A)*	Intron 1	Higher inducibility	rs762551	49.6	61.3[a]	30.5[a]	50.7	na	na
**2 (63C>G)*	Exon 2	Phe21leu	rs56160784	na	na	na	0	na	na
CYP2A6									
**1B*	3′ UTR	Increased		11.2[a]	42.8[a]	27.6	32.7	na	na
**2 (479T>A), (1799T>A)*	Exon 3	Leu160His	rs1801272	0.3	0	2.3	na	1.0	na
**4A*	–	Gene deletion	–	0.5[a,b]	11	3.0[a,b]	11.3	8.9	na
**5 (1436G>T), (6582G>T)*	3′ UTR, Exon 9	Gly479Val	rs5031017	0	0.5	0	na	0.7	na
CYP2C8									
**2 (805A>T)*	Exon 5	Ile269Phe	rs11572103	17[b]	0	0.3	3.0	0.8	na
**3 (416G>A), (1196A>G)*	Exon 8	Arg139Lys	rs10509681	0	0	10.9[b]	4.0	2.9	na
CYP2C9									
**2 (430C>T)*	Exon 3	Arg144Cys	rs1799853	2.8[a]	0	10.7[b]	9.0[b]	3.6[a]	na
**3 (1075A>C)*	Exon 7	Ile359Leu	rs1057910	2.0[a,b]	1.1[a,b]	7.4	9.7	8.0	na
**5 (1080C>G)*	Exon 7	Asp360Glu	rs28371686	2.2	0	0	0	na	na

TABLE 46.1 Continued

Gene and SNP	Location	Effect	dbSNP ID	Minor Allele Frequency					
				Africans	*Asians*	*Caucasians*	*NI*[*]	*SI*[*]	*NEI*[*]
CYP2C19									
**2 (681G>A)*	Exon 5	Splicing defect	rs4244285	16.0[a,b]	28.4[b]	13.3[a,b]	33.1	36.8	na
**3 (636G>A)*	Exon 4	Premature stop codon	rs4986893	0	10.1[a,b]	0.2[a,b]	1.9[b]	1.1	na
**17(−806C>T), (−340C>T)*	5′FR	Increased transcription	rs12248560	26.3	1.5[b]	19.1	na	19.2	na
CYP2D6									
**2(2850C>T), (4180G>C)*	Exon 2 and 6	Arg296Cys, Ser486Thr	rs16947, rs1135840	23.0[b]	14.0[a,b]	4.0[a,b]	29.3	34.8	na
**3 (2549delA)*	Exon 5	Frame shift	rs35742686	0.4[a]	1.0[a]	2.8[a]	9.2	0	na
**4 (1846G>A)*	Intron3-Exon 4 junction	Splicing defect	rs3892027	6.8[a]	4.0[a,b,c]	14.0[c,b]	11.5[b]	7.3	8.7
**5*	–	Gene deletion		6.6[a,b]	7.0[a,b]	6.9[a,b]	1.9	1.8	na
**6 (1707delT)*		Frame shift	rs5030655	0	0	1.5	0	na	na
**9 (2615-2617 delAAG)*		Lys281del	rs5030656	0	0	2.0	0	na	na
**10 (100C>T)*	Exon 1	Pro34Ser	rs1065852	5.6[a,b]	39.4[a,b]	1.5[a,b]	27.2[b]	10.2[a]	na
**14 (1758G>A)*	Exon 3	Gly169Arg	rs5030865	na	1.5	na	na	0	na
**17 (1023C>T), (2850C>T)*	Exon 2	Thr107Ile, Arg296Cys	rs28371706, rs16947	20.9	0	0.1	0	0	na
**29 (1659G>A), (1661G > C), (3183G>A)*		Val136Ieu, Val338Met	rs61736512, rs59421388	7.2	0	0	0	na	na
**41 (2988G>A)*		Aberrant splicing	rs28371725	nd	1.9[a]	8.5	12.5	na	na
1xN		Gene Duplication	–	1.2	0.1	0	0.6	na	na
2xN (+2850C>T, 4180G> C)		Gene Duplication	–	1.6	0.5	1.5	1.6	na	na
4xN (+1846G>A)		Gene Duplication	–	0.3	0	0	0.3	na	na
10xN (+100C>T)		Gene Duplication	–	na	0.5	0	0	na	na
41xN (+2988G>A)		Gene Duplication	–	na	na	0	0	na	na

(Continued)

TABLE 46.1 Continued

Gene and SNP	Location	Effect	dbSNP ID	Minor Allele Frequency					
				Africans	*Asians*	*Caucasians*	*NI**	*SI**	*NEI**
CYP2E1									
**1B (9896C>G)*	Intron 7	–	rs2070676	65.9[a,b]	18.1[a,b]	13.5	13.0	14.2	na
**4 (4768G>A)*	Exon 4	Val179Ile	rs6413419	17.5	0.1	2.8	na	0	na
**5B (−1293G>C) (−1053C>T)*	Promoter	–	rs2031920 rs3813867	4.3[a,b,c]	14.0[a,b,c]	3.8[a,b,c]	8.4[b,c]	1.3[a]	0.8[a]
**6 (7632T>A)*	Intron 6	–	rs6413432	10.1[a,b,c]	19.6[c]	7.7[a,b,c]	17.7[c]	22.2[c]	0.8[a,b]
CYP3A4									
**1B (−392A>G)*	5′ UTR	–	rs2740574	26.0[a]	0.1[a]	7.0[a]	1.2	na	na
**2 (15713T>C)*	Exon 7	Ser222Pro	rs55785340	0	na	0	0	na	na
**4 (13871A>G)*	Exon 5	Ile118Val	rs55951658	na	2.4	na	0	na	na
**5 (15702C>G)*	Exon 7	Pro218Arg	rs55901263	na	0.7	na	0	na	na
**6(17661insA)*	Exon 9	Frame shift	rs4646438	na	0.5	na	0	na	na
**10 (14304G>C)*	Exon 6	Asp174His	rs4986908	2.4[+]	0[+]	0.6[+]	0	na	na
CYP3A5									
**2 (27289C>A)*	Exon 11	Thr398Asn	rs28365083	0[+]	0	1.0	na	0	na
**3 (6986A>G)*	Intron 3	Splicing defect	rs776746	15.0[a,b]	80.6[a,b]	91.6[a,b]	68.2[b]	56.0[a]	na
**4 (14665A>G)*	Exon 7	Gln200Arg	rs56411402	0[+]	0	0	na	0	na
**6 (14690G>A)*	Exon 7	Lys208Lys	rs10264272	14.2	0	0.1	0	0	na
−44A>G	Promoter	–		na	28.2	9.2[a]	34.7	na	na
GSTM1									
**0 (null)*	–	Gene deletion	–	36.6[a,b]	42.9[a,b,c]	57.0[a,b,c]	30.6	28.8[c]	34.4[b]
GSTT1									
**0 (null)*	–	Gene deletion	–	41.6[c,b]	49.2[c,b]	19.7	17.2	14.8[c]	19.7[b]
GSTP1									
**B (313A>G)*	Exon 5	Ile105Val	rs1695	36.8[a,b,c]	20.6[b]	25.9[b]	23.8[b]	14.6[a,c]	21.7[b]
UGT1A1									
**28 (TA)6>7*	Promoter	Abolished activity	rs8175347	55.3	13.1[b]	29.6	na	39.7	na
UGT1A7									
**2 (387T>G), (392G>A)*	Exon 1	Asp129Lys; Arg131Lys	–	38.6	20.0[a]	22.0[a]	35.6	na	na

TABLE 46.1 Continued

Gene and SNP	Location	Effect	dbSNP ID	Minor Allele Frequency					
				Africans	*Asians*	*Caucasians*	*NI*[*]	*SI*[*]	*NEI*[*]
**3 (387T>G), (392G>A), (622T>C)*	Exon 1	Asp129Lys; Arg131Lys; Arg208Trp	–	22.8[a]	25.0[a]	42.4	35.6	na	na
**4(622T>C)*	Exon 1	Arg208Trp	rs1126802	0.7[a]	0.5[a]	0.5[a]	3.2	na	na
**12 (−57T>G)*	Promoter	Decreased	rs7586110	na	na	38.8	37.0	na	na
TPMT									
**2 (238G>C)*	Exon 5	Ala80Pro	rs1800462	0	0	0.3	0	0.1	na
**3A (460G>A), (719A>G)*	Exon 7 and 10	Ala154Thr; Tyr240Cys	rs1800460 rs1142345	0.2	0.3	4.5[a]	0.4	0	na
**3B (460G>A)*	Exon 7	Ala154Thr	rs1800460	0	0	0	0	0.1	na
**3C (719A>G)*	Exon 10	Tyr240Cys	rs1142345	7.6[a,b]	0.9	0.4[a,b]	2.1	1.1	na
**8 (644G>A)*	Exon 10	Arg215His	rs56161402	1.6	0	0	na	0	na
SULT1A1									
**2 (638G>A)*	Exon 7	Arg213His	rs9282861	28.5	8.7[c,b]	41.5[c,b]	na	22.6	27.2
COMT									
472G>A	Exon 4	Val158Met	rs4680	34.0[a,b]	27.5[a,b]	48.5[b]	49.0[b]	41.6[a]	na
5289A>G	Intron 1	–	rs3788319	na	na	na	na	47.3	na
4239T>C	Intron 1	–	rs737865	10.6[b,+]	32.9[b,+]	32.6[b,+]	na	24.8	na
VS2 (98A>G)	Intron 2	–	rs6269	na	na	40.5	na	31.1	na
186C>G	Exon 4	Leu136Leu	rs4818	na	na	39.8	na	32.0	na
36C>T	Exon 3	His62His	rs4633	26.7[b,+]	23.3[b,+]	51.5[b]	na	44.0	na
522G>A	3′ UTR	–	rs165599	na	na	36.9	na	42.1	na
NAT2									
**5 (341C>T)*	Exon 2	Ile114Thr	rs56935242	47.2[b]	6[b]	43[b]	na	30.0	na
**6 (590G>A)*	Exon 2	Arg197Gln	rs60190029	28.7[b]	24.0[a,b]	29.0[b]	31.5	37.0	26.0[b]
**7 (857G>A)*	Exon 2	Gly286Glu	rs1799931	3.1[a,b]	15.2[b]	3.0[a,b]	13.3[b]	25.0[a]	11.1[b]
**11(481C>T)*	Exon 2	Leu161Leu	rs1799929	45.7[a,b]	5.1[a,b]	46.0[a,b]	29.2[b]	22.0[a]	na
**12 (803A>G)*	Exon 2	Arg268Lys	rs56599719 rs58999469	55.1[a,b]	9.8[a,b]	41.0[a,b]	25.0	29.0	na
**13 (282C>T)*	Exon 2	Tyr94Tyr	rs1041983	33.4[b]	30.5[b]	32.0[b]	na	44.0	na
**14 (191G>A)*	Exon 2	Arg64Gln	rs1801279	3.1	0	1.3	na	0	na
MDR1									
3435C>T	Exon 26	Ile1145Ile	rs1045642	16.5[a,b]	38.4[a,b]	49.0[a,b]	53.6[b]	59.5[a]	na

(Continued)

TABLE 46.1 Continued

Gene and SNP	Location	Effect	dbSNP ID	Minor Allele Frequency					
				Africans	*Asians*	*Caucasians*	*NI*[*]	*SI*[*]	*NEI*[*]
2677 G>T/A	Exon 21	Ala893Ser Ala893Thr Ala893Pro	rs2032582	0.9[a,b]	58.5[a]	43.5[a,b]	42.4[b]	60.0[a]	na
1236C>T	Exon 12	Gly412Gly	rs1128503	15.0[a]	65.2[a]	41.0[a]	51.9	na	na
OCT1									
1022C>T	Exon 6	Pro341Leu	rs2282143	8.0	16.7[b]	2.0[b]	na	8.9	na
1222A>G	Exon 7	Met408Val	rs628031	73.5	74.0	57.4[b]	na	80.3	na
1386C>A	Intron	Unknown	rs622342	15.6[b,+]	20.6[+]	41.0[b,+]	na	24.5	na
SLCO1B1									
**1B (388A>G)*	Exon 4	Asn130Asp	rs2306283	87.0[a]	64.0[a]	37.0[a]	45.0	nd	nd
**4 (463C>A)*	Exon 4	Pro155Thr	rs11045819	6.2[a,+]	1.1[+]	15.0[a,+]	2.6	nd	nd
**5 (521T>C)*	Exon 6	Val174Ala	rs4149056	9.0[a,+]	11.0[a,+]	15.8[a,+]	1.4	nd	nd
CYP19A1									
−588 G>A	5′FR exon 1.1	–	rs7176005	–	19.0[b]	8.8[b]	nd	35.9	nd
−144C>T	5′FR exon 1.1	–	rs6493497	28.3[b]	31.0	8.8[b]	nd	35.9	nd
IVS1G>A	Intron 2	–	rs749292	50.8[b]	63.3[b]	45.3[b]	nd	24.9	nd
	Intron	–	rs1062033	7.2[b]	54.4[b]	50.9[b]	nd	15.3	nd
−196T>G	5′FR exon 1.6	–	rs10459592	27.3	49.0[b]	60.3[b]	nd	33.7	nd
−77C>T	5′FR exon 1.6	–	rs4775936	5.9[b]	49.0[b]	53.0[b]	nd	19.9	nd
IVS2T>G	Intron 2	–	rs727479	22.9	24.4	32.0	nd	25.8	nd
240A>G	Exon 3	Val80Val	rs700518	13.1	48.0[b]	58.2[b]	nd	17.8	nd
790C>T	Exon 7	Arg264Cys	rs700519	21.4[b]	16.0[b]	2.7[b]	nd	33.7	nd
1531C>T	3′ UTR	–	rs10046	16.8	56.0[a,b]	56.6[a,b]	25.0	20.0	nd
1673G>T	3′ UTR	–	rs4646	32.6[b]	26.7[b]	23.9[b]	36.0	41.1	nd
VKORC1									
−1639G>A	Promoter	–	rs9923231	43.8[a,b]	92.4[a,b]	30.1[a,b]	14.0	10.0	nd
1173C>T	Intron 1	–	rs9934438	10.2	91.5[b]	19.9	nd	9.3	nd
2255C>T	Intron	–	rs2359612	19.4[b]	87.9[b]	19.9[b]	nd	8.8	nd
3730G>A	3′ UTR	–	rs7294	43.2[b]	5.8[b]	36.3[b]	nd	83.6	nd

TABLE 46.1 Continued

Gene and SNP	Location	Effect	dbSNP ID	Minor Allele Frequency					
				Africans	*Asians*	*Caucasians*	*NI**	*SI**	*NEI**
MTHFR									
1298A>C	Exon 7	Glu429Ala	rs1801131	11.5	18.6	34.1[b,c]	20.7[b]	8.7[a]	17.0
677C>T	Exon 4	Ala222Val	rs1801133	9.3	36.0[a,b,c]	31.0[a,b,c]	10.1	2.4[c]	15.1[b]
ACE									
I/D	–	–	rs4646994	59.4[a,b,c]	31.7	61.2[a,b,c]	20.0[b,c]	44.0[a]	36.1[a]

**Pooled mean allele frequency. +HapMap data derived from NCBI dbSNP database. [a]NI vs. Africans, Asians, Caucasians, SI, NEI. [b]SI vs. Africans, Asians, Caucasians, NI, NEI. [c]NEI vs. Africans, Asians, Caucasians, NI, SI.*

DME—drug-metabolizing enzymes; NI—North Indians; SI—South Indians; NEI—North East Indians; AA—aminoacid; nd—data not available.

Adapted from "Distribution of genetic polymorphisms of genes encoding drug metabolizing enzymes & drug transporters - a review with Indian perspective" by Umamaheswaran et al. 2014, Indian J Med Res [11].

therapeutic drugs including caffeine, clozapine, olanzapine, amitriptyline, R-warfarin, verapamil, theophylline, propranolol, clomipramine, imipramine, haloperidol, and acetaminophen. Further, the association of *CYP1A2* polymorphisms with increased risk of various cancers has been reported. The gene encoding CYP1A2 is located on chromosome 15q22 with 6 exons, extending 7.8 kilobases (kb).

In total, more than 150 different polymorphisms have been reported in *CYP1A2*. Nevertheless, the functional significance has been described only for *CYP1A2*1C (−3860G>A)* and *CYP1A2*1F (−163C>A)* polymorphisms. The only study that determined the frequency of *CYP1A2* in any of the Indian population was carried out in NI population with a frequency of about 8%, 24.5%, 9.5%, 50.7% and 0% respectively for *CYP1A2*1C, *1D, *1E, *1F* and **2* alleles [11]. Comparison between NI and other major populations exhibited significant interethnic differences.

46.3.1.3 CYP2A6

CYP2A6, a phase I drug metabolizing enzyme (DME) which metabolizes about 3% of pharmaceutical agents including letrozole, tegafur, coumarin, valproic acid, methoxyflurane, artesunate, disulfiram, halothane and fadrozole. Additionally, it also metabolizes nicotine, retinoic acids, environmental toxins, procarcinogens and steroids. The *CYP2A6* gene is mapped on chromosome 19q13.2 along with other subfamily members *CYP2A7* and *CYP2A13*. It consists of 9 exons and spans about 6 kb. Currently, 81 variant alleles are known for *CYP2A6* gene. The activity of *CYP2A6* is found to be influenced by the presence of genetic polymorphisms which may affect the efficacy and toxicity profiles of CYP2A6 substrates. Further, polymorphisms of *CYP2A6* have been linked to tobacco-related cancer risks. A wide interethnic variation exists in the frequency distribution of *CYP2A6* alleles.

Three *CYP2A6* functional alleles which result in abolished activity (**2, *4A*, and **5*) and a normal allele designated as **1B* have been well known. Of the above *CYP2A6* variants, **1B* is the most prevalent allele in Indians followed by **4*, whereas **2* and **5* are found to be rare. The polymorphic allele frequency of *CYP2A6*2* and **5* are available only for SI; similarly, for *CYP2A6*1B* is available only in NI population. The interethnic comparison between NI (32.7%) and other populations indicates similarity with that in Caucasians (27.6%) and significant difference with that in Africans (11.2%) and Asians (42.8%). The defective allele *CYP2A6*4A* is prevalent in Indians (NI, 11.3%, SI, 8.9%) and Asians (11%) but significantly lower than in Africans (0.5%) and Caucasians (3%). In contrast, the other variant *CYP2A6*2* allele is present at low frequency in SI (1%), Africans (0.3%), and Caucasians (2.3%), but is absent in Asians. In the same way, *CYP2A6*5* is lower in SI (0.7%), Asians (0.5%) and is completely absent in Africans and Caucasians.

46.3.1.4 CYP2C8

CYP2C8 is one of the most important members of the CYP2C subfamily and its protein sequence is different from other members. CYP2C8 accounts for the metabolism of approximately 5% of drugs used in routine clinical practice. It includes some antidiabetics (troglitazone, pioglitazone, rosiglitazone & repaglinide), antihypertensive (verapamil), antimalarial (chloroquine, amodiaquine), anticancer (paclitaxel), non-steroidal anti-inflammatory drugs (NSAID) (ibuprofen), 3-hydroxy-3-methylglutaryl-coenzyme (HMG-CoA) reductase inhibitor (cerivastatin), and antiarrhythmic (amiodarone) drugs; and endogenous compounds such as retinoic and arachidonic acid.

The *CYP2C8* gene is localized in the long arm of chromosome 10 and composed of 9 exons, spanning 31 kb. To date, 20 different variant alleles have been identified in the *CYP2C8* gene. The most widely studied variants responsible for reduced enzyme activity are *CYP2C8*2*, *CYP2C8*3, CYP2C8*4* and *CYP2C8*5*. Previous studies have reported that impaired metabolism of CYP2C8 substrates are associated with the presence of genetic polymorphisms in CYP2C8. Subjects carrying the homozygous genotypes (**2/*2* or **3/*3*) of *CYP2C8*2* and *CYP2C8*3* polymorphisms display greater AUC values and decreased drug clearance. In addition, relationship between *CYP2C8* polymorphisms and risk of myocardial infarction and gastrointestinal bleeding has been observed.

The only study that determined the frequency distribution of *CYP2C8* alleles in any of the Indian population was carried out in SI Tamilians with a frequency of about 0.8% and 2.9% for *CYP2C8*2* and **3*, respectively. Studies in different populations from around the world specify that *CYP2C8*2* is rare allele in SI and Caucasians (0.3%), whereas it is a common allele in Africans (17%). *CYP2C8*3* variant is common in Caucasians (10.9%) while it occurs at a small frequency in Indians (NI, 4%; SI, 2.9%), and is absent in Africans. Regarding Asians, the *CYP2C8* polymorphism is monomorphic for *CYP2C8*1* allele [11].

46.3.1.5 CYP2C9

CYP2C9 is a predominant member of CYP2C isoforms and constitutes about 20% of the CYP450 liver content. It participates in the oxidative metabolism of the following drugs: S-warfarin, rosiglitazone, tolbutamide, phenytoin, glyburide, glibenclamide, glimepiride, glipizide, losartan, irbesartan, torsemide, tamoxifen, fluvastatin, fluoxetine, amitriptyline and other commonly used anti-inflammatory drugs (diclofenac, ibuprofen, naproxen, piroxicam, aceclofenac, celecoxib). The *CYP2C9* gene is highly polymorphic and is located on chromosome 10q24.2.

Up to now, 41 genetic variants have been identified for the *CYP2C9* gene in different populations. The most common being *CYP2C9*2 (430C>T)* in exon 3, and *CYP2C9*3 (1075A>C)* in exon 7. These non-functional alleles have been described as a risk factor for various diseases such as MI, hypertension, colorectal cancer and major depressive disorders; also with certain life threatening adverse drug reactions including gingival hyperplasia, hypoglycaemia and gastrointestinal bleeding. Based on CYP2C9 metabolic activity, individual phenotypes are divided into extensive metabolizers (EMs), Intermediate metabolizers (IMs) and poor metabolizers (PMs).

The frequency of *CYP2C9*2* is higher in NI (9%) than that in SI (3.6%), however the *CYP2C9*3* (NI, 9.7%; SI, 8%) allele is equally distributed among the Indian populations (see Table 46.1). The distribution of *CYP2C9*2* in NI is in agreement with that in Caucasians (10.7%) but showed difference with that in Africans (2.8%). In contrast, the SI showed similarity with Africans but was different from that in Caucasians. The prevalence of *CYP2C9*3* is higher in Indians than that described in Africans (2%) and Asians (1.1%) but no difference was observed with Caucasians (7.4%).

46.3.1.6 CYP2C19

CYP2C19 is an essential member of the CYP450 superfamily and it contributes about 16% of total hepatic content. CYP2C19 is the principal enzyme involved in the hepatic metabolism of drugs such as antimalarial (proguanil), oral anticoagulants (R-warfarin), chemotherapeutic agents (cyclophosphamide), anti-epileptics (S-mephenytoin, diazepam, phenobarbitone), antiplatelets (clopidogrel), proton pump inhibitors (omeprazole, pantoprazole, lansoprazole, rabeprazole), antivirals (nelfinavir), and antidepressants (amitriptyline, clomipramine). The gene encoding CYP2C19 is located on chromosome 10q24 and currently 35 polymorphisms are known in the CYP2C19 gene. The variant genotypes of *CYP2C19*2 (681G>A), CYP2C19*3 (636G>A)* and *CYP2C19*17 (−806C>T & −340C>T)* are major factors attributed to interindividual differences in the pharmacokinetics and response to *CYP2C19* substrates. *CYP2C19*2* and **3* are associated with diminished enzyme activity, whereas *CYP2C19*17* results in increased activity.

Notably, the *CYP2C19*2* allele occurs at a higher frequency among Indian populations (NI, 33.1%; SI, 36.8%) than among Africans (16%), Caucasians (13.3%) and Asians (28.4%). The *CYP2C19*3* allele in Indians is 1.9% in NI and 1.1% in SI. The distribution of loss-of-function alleles (*CYP2C19*2* and **3*) show greater inter- and intraethnic variations among Indian populations compared to other major ethnics such as Asians (28.4% and 10.1%) and Caucasians (13.3% and 0.2%) individually. In contrast, *CYP2C19*17* resulting in ultra-rapid metabolism of *CYP2C19* substrates, studied only in SI Tamilians, showed a frequency of about 19.2%. The comparison of SI with Asian (1.5%) subjects indicates significant difference and matches with that in Africans (26.3%) and Caucasians (19.1%) for *CYP2C19*17* allele (see Table 46.1).

46.3.1.7 CYP2D6

Debrisoquine hydroxylase (CYP2D6), a major hepatic enzyme is encoded by the *CYP2D6* gene in humans. It plays a pivotal role in the metabolism and elimination of at least 25% of commonly prescribed drugs including opioids, antiestrogen, antiarrhythmic, antipsychotics, antidepressants, β-blockers, antihypertensives and antihistamines. The *CYP2D6* gene has polymorphic expression and is the most widely investigated gene in relation to interindividual and interethnic variability in the treatment outcome of CYP2D6 substrates. The *CYP2D6* gene locus has been mapped to 22q13.2 and composed of 9 exons with 8 introns. To date, 135 polymorphic variants have been reported in the CYP2D6 gene sequence and their frequencies differ markedly among the global populations. The most frequent polymorphisms are *CYP2D6*2, *3, *4, *5, *10, *14, *17* and **41*. Depending upon the *CYP2D6* allele combination, individuals phenotypes are classified into PMs, IMs and UMs.

The alleles associated with diminished CYP2D6 activity (**9, *17* and **29*) as well as the loss-of-function allele **14* were not detected in Indian populations. Of the several genetic variants, *CYP2D6*2* is the most prevalent functional allele in Indians, found at comparable frequencies between NI (29.3%) and SI (34.8%). SI showed significant difference compared with that in Asians (14%) and Caucasians (4%). On the other hand, NI showed similarity with Asians, but the differences are significant with Caucasians. The *CYP2D6*3* allele was found only in NI (9.2%) and its frequency is significantly higher as compared to that in Africans (0.4%), Asians (1%) and Caucasians (2.8%). The frequency of *CYP2D6*5* polymorphism, which lead to gene deletion was determined to be about 1.9% in NI and 1.8% in SI. A significantly lower *CYP2D6*5* frequency was observed in Africans (6.6%), Asians (7%) and Caucasians (6.9%).

The distribution of *CYP2D6*4* is higher in NI (11.5%) and Caucasians (14%), and is lower in Asians (4%) than those described in SI (7.3%), NEI (8.7%) and in Africans (6.8%). *CYP2D6*10* frequency difference is statistically significant between NI (27.2%) and SI (10.2%). The Indian populations showed significant difference compared with other populations with regard to *CYP2D6*10* (see Table 46.1). The *CYP2D6*41* has high allele frequency in NI (12.5%) and Caucasians (8.5%) and is low in Asians (1.9%).

The *CYP2D6* alleles responsible for gene duplication were studied only in NI Gujarati population. Overall, the ultra-rapid metabolism (*1xN* and *2xN*) and loss of enzyme activity (*4xN*) alleles were found at lesser frequencies in world populations, but abridged activity alleles (*10xN* and *41xN*) were found to be absent in NI [11].

46.3.1.8 CYP2E1

Human CYP2E1 is an N-nitrosodimethyl-amine demethylase, and belongs to the CYP450 super family. It is highly expressed in liver and the levels elevate in pathophysiological conditions such as fasting, diabetes, obesity and alcohol consumption. The *CYP2E1* gene is localized to chromosome 10q26.3, consists of 9 exons and 8 introns. The expressed enzyme catalyzes the biotransformation of several low molecular weight procarcinogens and protoxins, which includes N-nitrosamines, aniline, vinyl chloride, benzene, styrene, and urethane. In addition, it is involved in the metabolism of drugs such as acetaminophen, isoniazid, chlorzoxazone and fluorinated anesthetics.

The polymorphism *CYP2E1*1B* in intron 7, *CYP2E1*5B* in the 5'flanking region and *CYP2E1*6* in intron 6 have drawn much attention because of their association with inter-individual variation, altered gene function and expression. Many studies have reported the association of CYP2E1 genetic variations with liver disease and cancers. *CYP2E1*1B* frequency is evenly distributed in NI (13%), SI (14.2%) and Caucasians (13.5%), while it is significantly higher in Asians (18.1%) and even higher in Africans (65.9%). With regard to *CYP2E1*5B*, similarity was observed between SI (1.3%) and NEI (0.8%) but differences were seen in NI (8.4%). Likewise, there was observed similarity between NI and SI (17.7% vs. 22.2%) but differences in NEI (0.8%) with respect to *CYP2E1*6* (see Table 46.1). The frequency of *CYP2E1*5B* in NI is different from that in other ethnic populations whereas CYP2E1*6 is similar to that in Asians (19.6%) but different from that in Africans (10.1%) and Caucasians (7.7%). The SI are different from other major populations for *CYP2E1*5B* but matches with that in Asians (14%); for *CYP2E1*6* SI are different from that in Africans and Caucasians. The frequency of *CYP2E1*5B* and **6* alleles in NEI as compared to that in other ethnicities indicates statistically significant difference.

46.3.1.9 CYP3A4

The human cytochrome P450 3A4 (CYP3A4) is the largest member of the CYP3A subfamily and accounts for 30–60% of the total for CYP450 adult liver. The *CYP3A4* gene is localized on chromosome 7q21 and up to now, 41 CYP3A4 alleles have been identified. Among them *CYP3A4*1B (G>A)* and **22 (C>T)* are the defining alleles that have been reported. Genetic variations in *CYP3A4* are a major source for interpatient variability in plasma concentration, adverse events and therapeutic response to drugs such as paclitaxel, fentanyl, tamoxifen, tacrolimus, and statin. In addition, existing studies have reported the role of *CYP3A4* defective alleles in the disease predisposition to prostate cancer, estrogen receptor-negative breast cancer and type 2 diabetes. The prevalence of *CYP3A4*1B* is known in NI (1.2%) only and the comparison between NI and other populations indicates significant variations. The other *CYP3A4* variants such as **2, *4, *5, *6* and **10* were not detected in NI [11].

46.3.1.10 CYP3A5

Drugs undergoing oxidative metabolism including immunomodulators (tacrolimus, cyclosporine), antipsychotics (olanzapine), antiestrogen (tamoxifen), anticancer (irinotecan, docetaxel, vincristine), antimalarial (mefloquine, artemether, lumefantrine), antihistamines (chlorpheniramine, terfenadine, astemizole), antiplatelets (clopidogrel), antihypertensives (nifedipine, amlodipine, felodipine, verapamil), antivirals (indinavir, nelfinavir, ritonavir, saquinavir), antibiotics (clarithromycin), HMG-CoA reductase inhibitors (atorvastatin, cerivastatin, lovastatin) and steroids (testosterone, estradiol, progesterone and androstenedione), are mainly metabolized by CYP3A5 isoenzyme. Its aminoacid sequence is identical to CYP3A4.

The gene *CYP3A5* is located adjacent to *CYP3A4* on chromosome 7. It contains four major polymorphisms, *CYP3A5*2 (27289C>A), *3 (6986A>G), *4 (14665A>G),* and **6 (14690G>A)*. Out of them, *CYP3A5*3* and **6* alleles result in absence of enzyme activity. *CYP3A5* genetic polymorphisms cause variation in its enzyme activity, which markedly influences the pharmacokinetics of *CYP3A5* substrates. Further, individual susceptibility to hypertension, acute myeloid leukemia (AML) or ALL, and breast cancer may be increased by non-functional alleles of *CYP3A5*.

The only polymorphic allele determined in Indian population is *CYP3A5*3*. The other variants **2, *4* and **6* were found to be monomorphic for *CYP3A5*1*. Allele frequency of *CYP3A5*3* in SI (56%) is statistically significantly as compared to NI (68.2%). Likewise, Indian populations are statistically different compared with that in Caucasians (91.6%), Asians (80.6%) and Africans (15%). Further, a polymorphism *(−44A>G)* located in the promoter region of CYP3A5 was observed only in NI (34.7%) which frequency is higher than in Caucasians (9.2%) and similar to that in Asians (28.2%)

46.3.1.11 GSTM1, GSTT1, and GSTP1

Glutathione S-transferases (GSTs) are polymorphically expressed phase II drug metabolizing enzymes. They play a key role in the inactivation of a broad range of xenobiotics and endogenous substances and also protect cells from oxidative stress. The GST super family consists of 5 classes, namely alpha, mu, theta, pi, and gamma. The gene encoding the common members of GSTs are *GSTM1* mu (chr1p13.3), *GSTT1* theta (chr22q11.2), and *GSTP1* pi (chr11q13.2).

Among the 3 GSTs, *GSTM1* and *GSTT1* are the most polymorphic genes. The null genotypes *GSTM1 *0/*0* and *GSTT1 *0/*0* results in polymorphic gene deletions and are responsible for the complete absence of their respective enzymes. A coding region polymorphism in *GSTP1*, which substitute's isoleucine by valine at codon 105 is associated with diminished enzyme activity. Development of carcinogenesis in colon, bladder, lung, prostate and digestive tract are related to genetic polymorphisms in *GSTM1, GSTT1,* and *GSTP1* genes. Additionally, they have an impact on drug metabolism and thereby cause interpatient differences in disease prognosis, efficacy and toxicity of GST substrates.

GSTM1 and *GSTT1* null allele frequencies were found to be 30.6% and 16.7% in NI; 28.8% and 14.8% in SI and 34.4% and 19.7% in NEI (see Table 46.1). With regard to *GSTM1* and *GSTT1* null polymorphisms, a general uniformity was observed between NI *versus* SI and NI versus NEI. In contrast, uniformity is significantly dissimilar among SI and NEI. The frequency pattern of *GSTM1* in Indian population is statistically different from that in Africans, Asians and Caucasians. The frequency of *GSTT1* polymorphism in SI is similar with that in NI and different from that in NEI. Further, GSTT1 frequency in Indian population is significantly less as compared to that in Africans and Asians, but showed similarity with that in Caucasians (see Table 46.1). The *GSTP1 105Val* allele frequency is statistically significant in SI (14.6%), as compared to that in NI (23.8%), NEI (21.7%), Africans (36.8%), Asians (20.6%), and Caucasians (25.9%). A significant difference is also noted in Africans compared to NI and NEI.

46.3.1.12 UGT1A1

The uridine diphosphate glucuronosyl transferase 1A1 (UGT1A1) is an important member of the UGT1A family. It facilitates the metabolism and subsequent excretion of potentially toxic xenobiotics, carcinogens, estrogens, bilirubin and therapeutic drugs. The *UGT1A1* gene spans about 160 kb and is localized on chromosome 2q37. Polymorphisms in the *UGT1A1* gene confer low enzyme activity that eventually leads to interpatient differences in the pharmacokinetics of UGT1A1 substrates, particularly anticancer agent irinotecan. The *UGT1A1* gene is highly polymorphic, as of now, more than 50 polymorphisms have been reported, with the most characterized being *UGT1A1* (TA) 6>7 *(UGT1A1*28)* polymorphism in the 5'promoter region. This TATA box polymorphism is characterized by 7 thymidine-adenine (TA) repeats instead of 6, which results in lower promoter activity.

The only study to determine the frequency of *UGT1A1**28 in native Indians was carried out in SI. This study found *UGT1A1*28* frequency to be as follows: Andhrite (32.2%), Kannadiga (29%), Keralite (46.5%) and Tamilian (52.8%) (see Table 46.1). The highest prevalence of (TA) repeats was found in Keralites and Tamilians, which suggests that these two populations are more predisposed to development of cancer, Gilbert syndrome, Crigler-Najjar syndrome and severe toxicity when treated with irinotecan. Statistically significant differences in the frequencies of this variant were observed in SI (39.7%), Africans (55.3%), Asians (13.1%), and Caucasians (29.6%).

46.3.1.13 UGT1A7

Extra hepatic phase II DME, uridine 5′-diphosphate-glucuronosyltransferase 1A7 (UGT1A7) catalyzes the glucoronidation and detoxification of numerous endogenous and exogenous substances, such as bilirubin, drugs, steroid hormones, phenols, coumarin and environmental carcinogens. It is expressed in several tissues including lung, stomach, pancreas, oropharynx and esophagus. The human monooxygenase enzyme, UGT1A7 is encoded by the *UGT1A7* gene, which is located on chromosome 2q37. At present, 11 missense variants are known in the *UGT1A7* gene. Four of them are most common polymorphisms, *UGT1A7*1* (N129 R131 W208), *UGT1A7*2* (K129-K131-W208), *UGT1A7*3* (K129-K131-R208), and *UGT1A7*4* (N129-R131-R208).

Genetic susceptibility to cancers such as liver, oral, colon, gastrointestinal, and pancreas, as well as chronic pancreatitis are attributed to *UGT1A7* genetic variations, which modulates the enzyme's detoxification activity. *UGT1A7* gene polymorphism frequencies are available only in NI; they have not been studied in other Indian populations (see Table 46.1). On the whole, however, the *UGT1A7* polymorphic alleles seem to be more common in NI. Compared with those in Africans, Asians, and Caucasians significant interethnic differences have been observed for *UGT1A7*2, *3* and **4* alleles. Conversely, no such difference have been observed for *UGT1A7*12*.

46.3.1.14 TPMT

The gene thiopurine S-methyltransferse (*TPMT*) encodes the cytosolic enzyme TPMT which is mapped on chromosome 6p22.3 and has 10 exons spanning over 27 kb in size. TPMT, a rate limiting enzyme, plays a key role in the metabolism of thiopurine drugs (6-mercaptopurine, azathioprine and 6-thioguanine). Heritable deficiency of TPMT activity, regulated by genetic polymorphisms of the *TPMT* gene is an important factor causing interpatient variability in the clinical outcomes of thiopurine drugs. In erythrocytes, TPMT enzyme activity exhibit trimodal distribution in Caucasians. On the other hand, Southeast Asians show unimodal distribution.

Among the 24 genetic polymorphisms reported in *TPMT* gene, four common polymorphisms cause 80-95% intermediate and low enzyme activity. These are *TPMT*2 (238G>C),*3A (460G>A* and *719A>G),*3B (460G>A)* and **3C (719A>G)*. Patients who carry the defective alleles of TPMT are more likely to experience life threatening toxicity when they are treated with thiopurines. In Indians, *TPMT*3C* is the most common genetic variant; **2* and **3B* are scarce; and the African-specific allele, *TPMT*8* is absent.

*TPMT*3A* allele causing the largest decrease in TPMT activity, has been detected in NI (0.4%) but is absent in SI. Similarly, *TPMT*2* and *TPMT*3B* are present only in SI (0.1% and 0.1%) but not in NI (see Table 46.1). With regard to *TPMT*3A*, the NI (0.4%) are divergent from Caucasians (4.5%) and comparable to Asians and Africans. With respect to *TPMT*3C*, the NI (2.1%) and SI (1.1%) are different from that in Caucasians (0.4%), Africans (7.6%) and matches with that in Asians. On the whole, the Indians display higher TPMT activity due to less prevalence of TPMT alleles. Therefore, Indians are likely to be at lower risk of developing toxicity compared to Africans, Asians and Caucasians.

46.3.1.15 SULT1A1

Sulphotransferase 1A1 (SULT1A1) is the most widely expressed member of the sulphotransferase family. Through sulfation, it plays a significant role in the detoxification and bioactivation of several procarcinogens, toxins and xenobiotics. The catalytic activity and thermostability of SULTIA1 protein has been modified by the presence of genetic polymorphisms in the *SULT1A1* gene. Allelic variants and copy number variations of *SULT1A1* are implicated in the clinical efficacy and safety profiles of tamoxifen, as well as in susceptibility to hormone-dependent cancers.

The *SULT1A1* gene is extremely polymorphic and located on the short arm of chromosome 16; and is sized about 180,005 bases. Two coding variants *SULT1A1*2 (638G>A)* and *SULT1A1*3 (667A>G)* are the notable genetic polymorphisms; their frequency vary in ethnically diverse populations (see Table 46.1). The normative frequency of *SULT1A1*2* in Indians was established only in SI (22.6%) and NEI (27.2%) populations but not in NI. As shown in Table 46.1, *SULT1A1*2* frequency is significantly higher than those in Asians (8.7%) and lower than in Caucasians (41.5%), however, similar frequency is observed between SI, NEI and Africans (28.5%).

46.3.1.16 COMT

Catechol-O-methyltransferase (COMT), a phase II DME regulates the O-methylation of catecholamine neurotransmitters and catechol estrogen. The chromosome location of the gene encoding COMT protein is 22q11.2 which translates two different proteins, a low-affinity COMT (S-COMT) and high-affinity (MB-COMT) by alternative gene regulation. Until now, 30 different polymorphisms have been identified for the COMT gene. Among them, *COMT rs4680* (*472G>A*) in exon 4 is the most characterized functional polymorphism. This polymorphism contributes about 3 to 4 fold reduced methylation activity.

Individuals possessing the polymorphisms are susceptible to breast cancer, endometrial cancer, Parkinson's disease, and schizophrenia; and exhibit interindividual variation in pain sensitivity and therapeutic response to drugs used for chronic pain, including analgesics. The frequency of *COMT* polymorphic alleles namely *rs3788319, rs737865, rs6269, rs4818, rs4633,* and *rs165599* are available only in SI, excluding the SNP *rs4680*. That of *rs4680* is significantly different between the Indian populations (NI, 49% vs. SI, 41.6%). Compared to NI, significant differences are seen in Africans (34%) and Asians (27.5%). The differences in *COMT rs4680, rs737865* and *rs4633* frequencies for SI are statistically significant compared to Africans, Asians and Caucasians [11].

46.3.1.17 NAT2

Arylamine N-acetyltransferase enzymes NAT1 and NAT2 are polymorphic drug metabolizing enzymes. They participate in the detoxification and metabolic activation of arylamines, arylhydroxylamines, arylhydrazine and other drugs. The acetylation activity of the enzyme can be classified into slow, intermediate and rapid which is determined by the individual's *NAT2* genotype. The genes encoding both isoforms are mapped on chromosome 8p21.3-23.1. The common functional polymorphic alleles of *NAT2* are *191G>A* (Arg64Gln), *282C>T* (Tyr94Tyr), *481C>T* (Leu161Leu), *341T>C* (Ile114Thr), *590G>A* (Arg197Gln), *803A>G* (Arg268Lys) and *857G>A* (Gly286Glu). The most studied of these alleles are *NAT2*5, NAT2*6* and *NAT2*7* at respective positions 341, 590 and 857.

The variant *NAT2*6* corresponding to an inactive enzyme has a frequency of 37% in SI, which matches with that in NI (31.5%) and differs from that in NEI (26%). Likewise, *NAT2*7* accounts for 13.3% in NI, 11.1% in NEI and 25% in SI. The other polymorphisms *NAT*5* and **13* were studied only in SI and found to occur with frequencies of 30% and 44%, respectively, while **14* was not detected at all. The frequency of NAT gene polymorphism **11* is different between NI and SI (29.2% *vs.* 22%), whereas **12* is evenly distributed. A comparison of *NAT2* allele frequencies in Indians with those in Africans, Asians and Caucasians reveals significant interethnic differences (see Table 46.1). Carriers of these alleles show variation in the metabolism of isoniazid, hydralazine, ribavirin, retigabine, sulphamethoxazole and may have greater susceptibility to develop neoplasias.

46.3.2 Drug-Transporter Genes

46.3.2.1 ABCB1

Adenosine triphosphate binding cassette (ABC) superfamily of transport proteins exports the toxic xenobiotics and cellular metabolites from intracellular to extracellular space. Among the 49 ABC transporters, P-glycoprotein (P-gp) is the major energy-dependent transmembrane efflux transporter. The 170kD P-gp is the product of multidrug resistant 1 (*MDR1* or *ABCB1*) gene which plays a significant role in drug resistance and drug disposition. Genetic variations of *ABCB1* have been implicated in the development of breast cancer, colorectal cancer, upper aerodigestive cancer, Parkinson's disease and ulcerative colitis.

P-gp has a wide range of drug substrates which includes clinically relevant agents - anti-neoplasmics (doxorubicin, actinomycin D, paclitaxel), analgesics (morphine), antibiotics (erythromycin, levofloxacin, sparfloxacin, rifampicin), antivirals (nelfinavir, indinavir, efavirenz), antiepileptics (phenytoin, phenobarbital), antidepressants (amitriptyline), antiarrthythmics (digoxin, verapamil), antilipidemic (atorvastatin), immunosuppressants (cyclosporin A, tacrolimus, rapamycin), antihypersentive (losartan), and steroids (aldosterone, cortisol, dexamethasone). Inter-individual variability in the pharmacokinetics and pharmacodynamics of these drugs are attributed to the presence of genetic polymorphisms in the *ABCB1* gene which can alter the activity and expression of P-gp.

The gene *ABCB1* has 29 exons ranging between 49 to 209 bp in size and localized at chromosome 7q21.1. To date, more than 50 ABCB1 polymorphisms have been reported. Among them, *2677G>T/A/C* in exon 21, *3435C>T* in exon 26, and *1236C>T* in exon 12 are the most extensively studied polymorphisms. The polymorphic allele frequencies of these polymorphisms are more common in Indian populations. Conversely, the frequency (51.9%) of *1236C>T* polymorphism is presented only in NI. A greater inter and intra-ethnic difference was observed between Indian and other major populations with respect to *2677G>T/A/C, 3435C>T* and *1236C>T* variant alleles [11].

46.3.2.2 SLC22A1

There are 3 OCT isoforms with similar membrane topology (OCT1, OCT2 and OCT3) in the SLC22 superfamily of transporters. Among them, OCT1 is mainly expressed in the liver and involved in the transport of cationic drugs such as metformin, levodopa and imatinib. The gene *SLC22A1*, encoding the protein OCT1 is mapped onto chromosome 6q25.3 with 11 exons, spanning 37 kb. Individuals carrying genetic variation in the *SLC22A1* gene show variation in the distribution and elimination of OCT1 substrates because of differential transporter activity.

As of now, the frequency distribution of *SLC22A1* polymorphisms has been documented only in SI Tamil population for the following common genetic polymorphisms *rs2282143, rs628031* and *rs622342*. Their frequency are 8.9%, 80.3% and 24.5% for respectively, *rs2282143 (T)*, *rs628031 (G)* and *rs622342 (C)*. Comparison of the above frequency showed significant difference with that in Caucasians. However, a similar frequency is observed in Africans (73.5%) and Asians (74%) for *rs628031*. In contrast, *rs622342* showed similarity with Asians (20.6%) but not with Africans (15.6%). In the same way, similarity was observed with Africans (8%) but not with Asians (16.7%) for the SNP *rs2282143*. Existing evidence suggest marked interethnic difference in the allele frequency of *SLC22A1* gene polymorphisms that can eventually lead to dissimilarity in the pharmacokinetics and clinical outcomes of cationic drugs which are substrates of OCT1.

46.3.2.3 SLCO1B1

The gene organic anion transporting polypeptide 1B1 (*OATP1B1*, also known as *SLCO1B1*) encodes the transmembrane protein OATP1B1, which transports anionic drugs such as statins, irinotecan, rifampin, repaglinide and methotrexate. The *OATP1B1* gene is located on chromosome 12p12.2 with 14 coding and 1 non-coding exon, and spans 108.59kb in size. Genetic polymorphisms in

OATP1B1 diminish the transporter function, a major factor influencing inter-individual differences in toxicity and therapeutic efficacy of *OATP1B1* substrates. In addition, prevailing studies have reported the association of *OATP1B1* polymorphisms with increased risk of statin-induced myopathy, methotrexate-induced gastrointestinal toxicity and gallstone disease.

The prevalence of the most common genetic variations *OATP1B1*1B, OATP1B1*4* and *OATP1B1*5* is available only in NI population. The frequency of these polymorphisms is reported to be 45%, 2.6% and 1.4% for *OATP1B1*1B*, **4* and **5*, respectively. A comparison of the variant allele frequency shows statistically significant differences with other populations (see Table 46.1). The frequency of OATP1B1*1B allele is higher in Africans (87%) and Asians (64%); lower in Caucasians (37%). Likewise, the variants **4* and **5* are higher in Africans (6.2% and 9%) and Caucasians (15% and 15.8%), but regarding **4* Asians (1.1% and 11%) show similarity with NI.

46.3.3 Drug-Target Genes

46.3.3.1 CYP19A1

The 58-kDa protein aromatase (CYP19A1) is the product of the *CYP19A1* gene and belongs to the cytochrome P450 superfamily, which facilitates the biotransformation of estrogens (E1 and E2) from androgens (androstenedione and testosterone). Genetic variants in the sequence of *CYP19A1* cause changes in enzyme activity and have been associated with sex hormone levels, etiology of estrogen-dependent disease, and clinical outcomes of third-generation aromatase inhibitors (anastrazole, letrozole, and exemestane). The frequencies of the minor alleles of the pharmacogenetic variants associated with efficacy and toxicity of AIs—*rs4646 (T), rs10046 (T), rs700519 (T), rs700518 (G), rs727479 (G), rs4775936 (T), rs10459592 (G), rs749292 (A), rs6493497 (T),* and *rs7176005 (A)*—are 41.1%, 20%, 33.7%, 17.8%, 25.8%, 19.9%, 33.7%, 24.9%, 35.9%, and 35.9%, respectively, in SI. The frequency of *CYP19A1* polymorphisms in NI is limited to *rs4646* (36%) and *rs10046* (25%) with a similar frequency. Comparison of the allele frequency between SI and other major populations [17] shows statistically significant greater interethnic difference.

46.3.3.2 VKORC1

Vitamin K epoxide reductase complex subunit 1, encoded by *VKORC1* (chromosome 16), the drug target for the anticoagulants warfarin and acenocoumarol is used to prevent and treat thromboembolic cases. These drugs are metabolized primarily by the enzyme CYP2C9. Genetic variations in the *CYP2C9* and *VKORC1* genes cause changes in enzyme activity and thereby affect the metabolism and gamma carboxylation of clotting factors II, VII, IX, and X, and of the proteins C and S, respectively. Wide interethnic variability exists in the frequency of *VKORC1* variants in various populations. The frequency of the promoter region polymorphism *(−1639G>A)* associated with a lower dose requirement of warfarin is similar in Indian populations (14% vs.10%) and lower than those reported in other major populations [13,16]. The risk of bleeding is less in Indian patients on anticoagulants compared to that in Africans, Asians, and Caucasians because of low frequency of *−1639G>A*. The other polymorphisms, *2255C>T* and *3730G>A*, show a significant difference in their minor allele frequency between SI and other populations, while the *1173C>T* polymorphism frequency is similar to Africans and Caucasians but different from Asians [17].

46.3.3.3 MTHFR

Polymorphisms in the gene encoding methylene tetrahydrafolate reductase (MTHFR) have been implicated in the pathogenesis of cancers, migraine, cardiovascular disorders, rheumatoid arthritis, and neural tube defects. In addition, they have been suggested as putative markers in predicting the therapeutic efficacy and toxicity of methotrexate. This enzyme metabolizes 5,10-methylene tetrahydrofolate into 5-methyl tetrahydrofolate and affects the activity of cellular cycles participating in various cellular mechanisms.

The gene *MTHFR* is located on chromosome 15 and the two polymorphisms, one in exon 4 *(677C>T)* and another in exon 7 *(1298A>C)*, are the most characterized of *MTHFR*'s genetic variants. The frequency of the thermolabile T variant of *677C>T* in Indian populations shows similarity with that in Africans and differs from that in Asians and Caucasians. Within Indian populations, SI (2.4%) and NEI (15.1%) frequency is significantly different but NI (10.1%) frequency is comparable [18–20]. The prevalence of the variant that replaces the amino acid alanine instead of glutamate at codon 429 is 8.7% in SI, which is significantly more than the prevalence reported in NI and Caucasian, but is in agreement with African, Asian, and NEI prevalences. NEI (17.0%) frequency is similar to that in Africans and Asians but different from that in Caucasians.

46.3.3.4 ACE

The angiotensin-converting enzyme (ACE) belongs to the renin angiotensin system pathway, which plays a key role in blood pressure regulation. The insertion/deletion *(ID)* polymorphism of *ACE* is characterized by insertion or deletion of 287 bp in intron 16 of the gene-encoding ACE located on chromosome 17q23.3. This polymorphism is associated with many diseases, including essential hypertension, myocardial infarction, cardiomyopathy, diabetes, renal disorders, cancers, and serum ACE levels. A significant difference in the allele frequencies of the *D* allele was found between SI [21] and NI [22] populations (SI, 44% vs.

NI, 20%). A similar difference was observed between NI and NEI (36.1%) but not between SI and NEI [20]. Moreover, Indian populations exhibit significant differences in *D* allele frequencies compared to Africans and Caucasians, but, those frequencies are same in Asians [23].

46.4 PHARMACOGENETIC STUDIES IN INDIA

Over the last 10 years, we have gained much information pertinent to the genetic factors that determine disease risk, variable response, and vulnerability to drug-induced adverse events in the Indian populations. The studies described here analyzed the pharmacogenetic variants in genes encoding proteins of ADME pathways and those involved in the metabolism of carcinogens and other environmental toxins. Nevertheless, most of them are disease association studies that attempted to elucidate the heritability of diseases including cancer, diabetes, cardiovascular, neurological, and psychiatric disorders. Investigations of the molecular aspects of pharmacotherapy are scarce. Some molecular epidemiological studies have documented a few significant findings that highlight the unique genetic realm of Indian populations. There are several examples, including:

- The 25-bp deletion in the myosin-binding protein-C3 (MYBPC3) that is associated with inherited cardiomyopathy risk is present only in Indians [24].
- The *POLG-CAG* repeat variation, which increases male infertility in other world populations, is not associated in Indian men with infertility [25].
- A disease risk candidate gene (*SLC11A1)* for tuberculosis, leprosy, and leishmaniasis is seen in other populations but not in Indians [26].

In addition, studies have found several population- or regional-specific disorders such as Handigodu disease, Madras motor neuron disease, and pseudo-cholinesterase deficiency among Vyshyas and others [27–29] Thus, most of the genetic variations found to be associated with diseases in other international populations are not associated in Indian populations because of their unique genetic diversity. PGx investigations on clinically important drugs in India examined a limited number of samples from SI and NI populations with the candidate gene (SNP) approach. Because of the panorama of genetic heterogeneity, the results from these studies cannot be extrapolated to the entire Indian population.

46.4.1 Anticancer Drugs

46.4.1.1 Tamoxifen

The selective estrogen receptor modulator tamoxifen is a well-established antiestrogen in the treatment and prevention of estrogen receptor–positive breast cancer. To exert the anticancer effect, tamoxifen, a prodrug, should be converted into its active metabolite endoxifen. Although this biotransformation is mediated by a series of drug-metabolizing enzymes, CYP2D6 is the primary enzyme involved. Women with decreased CYP2D6 activity (poor metabolizers) have reduced endoxifen levels that can lead to therapeutic failure resulting in cancer relapse [30]. We conducted a study to evaluate the impact of *CYP2D6* gene polymorphisms on breast cancer recurrence in 132 South Indian patients receiving adjuvant tamoxifen. Those with a reduced CYP2D6 activity score (≤ 0.5) were associated with increased risk of recurrence and shorter recurrence-free survival ($HR = 7.29$, $p < 0.001$) compared to an activity score of ≥ 1 [31]. We did not find any association of *CYP2C19* variant alleles (**2/*17*) with recurrence [32].

Recently, Kalyan Kumar and coworkers found significant correlation of a shorter survival period with *CYP2D6*4* and *SULT1A1*2* polymorphisms [33] in 140 patients with five-year follow-up. However, neither study found any relationship between toxicity and genetic polymorphisms and both examined only limited pharmacogenetic variants. Even though there are conflicting results in terms of *CYP2D6* genotype and tamoxifen outcomes, *CYP2D6* is still useful in predicting tamoxifen responsiveness and resistance. Still, pharmacogenetic assessment of the genes involved in endoxifen formation and the active metabolite's measurement may be useful in predicting tamoxifen's effect.

46.4.1.2 Letrozole

Letrozole, a third-generation aromatase inhibitor, is the principal drug used in the treatment of postmenopausal patients with both early- and advanced-stage endocrine-responsive breast cancer (BC). Response to and tolerability of letrozole are influenced by germline polymorphisms in the *CYP19A1* gene [34]. In an unpublished study conducted in our lab, we evaluated the effect of *CYP19A1* and *TCL1A* gene polymorphisms on the clinical outcomes of adjuvant letrozole in 199 hormone receptor positive (HR+) BC patients on letrozole (2.5 mg, OD). There was a significant difference in the proportions of allele and genotype frequencies for the rs4646 polymorphism between recurrence and nonrecurrence groups. The carriers of two copies of the variant allele "T" were shown to be at a 6.69-fold increased risk of recurrence.

Furthermore, we found that patients with the *TT* genotype have significantly shorter RFS (100.71 vs. 210.93 months; $p < 0.02$). However, the *TCL1A* gene polymorphism is not associated with patient-reported letrozole-induced musculoskeletal pain. This is the first study assessing the association of *CYP19A1* variants with letrozole efficacy and toxicity in the adjuvant setting. The outcome of the study could be seen as an initial step toward achieving personalized letrozole treatment in SI postmenopausal HR+BC patients.

46.4.1.3 Anthracyclines

Treatment with a anthracycline-based neoadjuvant chemotherapy (NACT) regimen improves the clinical outcome of locally advanced breast cancer. Nevertheless, its effectiveness is limited by chemo resistance, resulting in marked interpatient variability and cardiotoxicity, which is genetically determined by polymorphic variants in genes coding drug-metabolizing and transporter proteins. SI patients carrying a homozygous mutant *(TT)* of *MDR1* have been shown to have 2.26 times better response to NACT compared to those with the wild genotype *(CC)*, but this difference is not statistically significant [35].

Subsequent study by Kumar and colleagues demonstrated that, according to WHO and RECIST criteria, patients with high *MDR1* expression have respectively 3.6 and 14.4 times greater chance of responding poorly to NACT [36]. The NACT response rate is better in NI BC patients with *GSTM1* (OR = 2.06) and *GSTT1* (OR = 1.45) null genotypes, but no statistical significance was attained [37]. Subjects heterozygous for the *MDR1 1236C>T* polymorphism were associated with NACT response. Regarding 2–4 grade toxicity, the *C* allele of *MDR1 1236C>T* was associated with NACT response, but no such association was described for the other polymorphisms *(2677G>T/A* and *3435C>T)* [38]. In the gene–gene interaction model, *GSTM1* and *GSTP1 Ile105Val*, *GSTM1* and *GSTT1*, and *GSTT1* and *GSTP1 Ile105Val* display significant association with response as well as toxicity [39].

46.4.1.4 Thiopurines

Thiopurine drugs—6-mercaptopurine (6-MP), 6-thioguanine, and azathioprine—are used in the treatment of leukemia, autoimmune disorders, and solid tumors, as well as in organ transplantations. Among these purine and pyrimidine metabolites, 6-MP is the most prescribed drug for acute lymphoblastic leukemia (ALL). In one study, epistatic interactions between the polymorphisms of *TPMT (*3C*, **12)* and *ITPA (ex2*, *ex3)* showed association with 6-MP-induced toxicity in 90 SI ALL children. In addition, the single-locus ITPA *94 C>A* showed a significant effect on toxicity [40]. The researchers, in subsequent investigations, demonstrated the synergistic interactions between genes involved in thiopurine and folate pathways (*GCPII C1561T*, *RFC1 G80A*, *cSHMT C1420T*, *TYMS 5′-UTR 2R3R*, *TYMS 3′-UTR ins6/del6*, *MTHFR C677T*, *MTR A2756G*) and cumulatively determined 6-MP–mediated toxicity in South Indian ALL children [41]. No association between TPMT gene polymorphisms and 6-MP–induced toxicities was observed in NI ALL children, but the maintenance dose therapy of 6-MP is 31% lower in TPMT heterozygous mutants [42].

46.4.2 Cardiovascular Drugs

46.4.2.1 Clopidogrel

Antiplatelet therapy with clopidogrel and aspirin is often used in the prevention of cardio, peripheral, and cerebrovascular events. Clopidogrel is oxidized into the active metabolite (thiol) by several hepatic CYP450 enzymes, which in turn inhibit ADP binding to the P2Y12 receptor on the surface of the platelet and is eliminated by the MDR1 transporter. The mechanism underlying clopidogrel resistance is attributed to the polymorphic variants in the genes encoding these proteins. A study involving 149 SI Tamil patients presenting with ischemic heart disease on clopidogrel maintenance therapy (75mg/DD) reported that carriers of the loss-of-function alleles of *CYP2C19* (**2* and **3*) and *MDR1 (T)* have significantly reduced response to clopidogrel because of higher impedence values. Conversely, residual platelet activity is lower among ultrarapid metabolizers *(CYP2C19*17)*, but the researchers observed no significance [43,44].

Another study, working with the previous cohort, found a 2.78- and 2.45-fold greater risk of developing clopidogrel resistance in subjects with *CYP3A5*3* homo- and heteromutants. Similarly, carriers of the defective alleles *CT* and *CC* genotypes of *P2Y12* were found to have a 2.2-fold greater risk [45]. Both studies failed to measure the active metabolite, thiol. On the contrary, clopidogrel resistance was not associated with variants of *P2Y12* and *P2Y1* in 100 NI coronary artery disease (CAD) patients [46]. However, Shalia and colleagues revealed a trend toward poor clopidogrel response with the variant genotypes of *CYP2C19*2* and *iT744C* of *P2Y12* genes compared to their respective wild-type genotype at 24 hours [47].

46.4.2.2 Warfarin

Warfarin is the most widely prescribed anticoagulant for the prevention of thromboembolic events. Because it is a narrow therapeutic index drug, routine therapeutic drug monitoring (international normalized ratio, INR) is required for dose adjustments to avoid adverse drug reactions. Warfarin's dose requirements exhibit interindividual variability and are influenced by both genetic and nongenetic factors. Pavani and coworkers developed an algorithm using a multiple linear regression model that includes Vitamin K intake, *CYP2C9* (**2* and **3*), and *VKORC1* (**3*, **4*, *D36Y* and *−1639 G>A)*. Using this algorithm, they explained the 44.9% variability in warfarin dose requirements [48]. Their following extended investigation explained the 61% variability [49].

Very recently, a new pharmacogenetic algorithm was developed to predict maintenance and starting warfarin indose in SI patients. This study explained a similar variability of ~62.1% in dosing requirements. The investigators reported that *CYP2C9*1/*2*, **1/*3*, and **2/*3* variant genotype carriers require a lower daily maintenance dose—2.8 mg, 2.3 mg, and 2.2 mg, respectively—than do carriers of the *CYP2C9*1/*1* genotype (5.2mg). The *CYP2C9*, *VKORC1*, and *GGCX* gene polymorphisms were associated with reduced dose, excluding the polymorphisms *rs7196161, rs7294,* and *rs2108622,* which were associated with an increased dose [50]. In 145 NI individuals, *VKORC1-1639 AA* and *CYP2C9 *3/*3* homozygous mutants were associated with risk of over anticoagulation and warfarin-related adverse events [51]. A significant effect of *CYP2C9* (**2* and **3*) and *VKORC1 (C1173T* and *G-1639A)* genotypes on warfarin dose was observed by Shalia and colleagues in 83 NI patients operated on for aortic/mitral valve replacement [52].

46.4.2.3 Acenocoumarol

Acenocoumorol, a coumarin derivative, is the counterpart of warfarin in the prevention and treatment of thromboembolic events. Its dosage requirements are affected by genetic variants and clinical parameters. Rathore and colleagues were the first to generate a population-specific algorithm in NI patients and to explain the 41.4% dosage variation. A multiple regression algorithm showed superiority over clinical data in terms of sensitivity, specificity, rate of overestimation, rate of underestimation, overall accuracy, accuracy in drug-sensitive cases, and accuracy in drug-resistant cases. The *VKORC1* and *CYP4F2* polymorphisms explained the 37% acenocoumarol dosage variability and turned out to be the main genetic determinants [53]. A trend toward lower maintenance doses of acenocoumarol was detected in NI patients on mechanical heart valve replacement who carry the *CYP2C9* and *VKORC1* polymorphisms [54].

A pilot study attempted to estimate coumarin sensitivity using *CYP2C9* and *VKORC1* gene polymorphisms in patients carrying the Factor V Leiden mutation. It concluded that genotyping of *CYP2C9* (**2* and **3*) and *VKORC1* (*-16539G>A*) polymorphisms before coumarin prescription is essential in inherited thrombophilia [55]. Attempts to quantify the influence of genetic and demographic factors on acenocoumarol maintenance dose were carried out in 170 SI patients who underwent heart valve replacement. Patients with at least one variant allele of *CYP2C9* (**1/*2* or **1/*3*) and *VKORC1 (GA)* required a lower maintenance dose compared to homozygous wild genotype patients. Demographic and genetic variables together predicted 30.4% of the acenocoumarol dose requirement. *VKORC1 (-16539G>A)* came out as the principal predictor of acenocoumarol dose maintenance [56]. The data from the Indian studies, irrespective of ethnicity, emphasize the importance of *CYP2C9* and *VKORC1* polymorphisms in determining the therapeutic dosage of anticoagulants (warfarin and acenocoumarol).

46.4.2.4 Statins

Statins are inhibitors of 3-Hydroxy-3-methylglutaryl coenzyme A reductase, the most widely prescribed cholesterol-lowering medication in the treatment of conditions associated with cardiovascular abnormalities. Despite the imperative role of statins in cardiovascular protection, some patients show degrees of intolerance and substantial differences in response. This variability is thought to be modulated by genetic factors in the proteins involved in the metabolic pathway of statins along with demographic and environmental factors.

One study explored the association of 18 polymorphisms (six genes) with atorvastatin response in 265 newly diagnosed NI CAD patients. It concluded that the polymorphisms *ABCB1 (-41A>G)*, *HMGCR SNP29 G>T*, *rs5908A>G*, *rs12916C>T*, and *CYP7A1-204A>C* were significantly associated with diminished low-density lipoprotein-C reduction. On the other hand, mutants of *CETP TaqI, −629C>A*, and *APOAI PstI* polymorphisms were associated with a greater increase in high-density lipoprotein cholesterol. Additionally, compared to single polymorphisms, three-way gene–gene interaction modules (*CYP7A1 rs892871AA*, *APOAI PstIP1P1*, and *HMGCR rs12916CT*) were found to be better predictors for LDL-C–lowering levels [57]. No association was observed with interindividual response to atorvastatin and *ABCG8 D19H* in 213 CAD NI patients [58].

46.4.3 Oral Hypoglycemic Drugs

46.4.3.1 Metformin

Metformin, an insulin-sensitizing biguanide agent, is the primary medication of choice in the treatment of T2DM; it scores higher than the other classes of oral antidiabetic drugs because of its additional allied beneficial effects. However, the glycemic response of metformin is highly variable. It has been suggested that the clinical consequences of metformin are altered by the molecular variants of the gene encoding organic cation transporter 1 (OCT1), with conflicting results. In the Indian scenario, Saranya and coworkers found significant association of the *OCT1* and *ATM* genes with metformin response in 165 SI diabetic patients. The homozygous mutant *(CC)* genotype of *OCT1 rs622342* is significantly higher in nonresponders than in responders (25% vs. 2.7%). Similarly, the homozygous wild genotype (AA) is higher in responders than in nonresponders (61.9% vs. 42.3%). On the other hand, the *C* allele of *ATM rs11212617* was found to be higher among responders than nonresponders (40% vs. 28.2%). No significant association was observed with regard to adverse side effects [59].

46.4.3.2 Glibenclamide

Glibenclamide belongs to the class of sulphonyl ureas that is used either as monotherpay or in combination with biguanides in the management of T2DM. Among the CYP450 enzymes involved in its metabolism, CYP2C9 is a major contributor. Carriers of the *CYP2C9*2/**3 variant alleles are likely to risk sulphonyl urea-induced hypoglycemia and variable glycemic response. Considering the role *CYP2C9* polymorphisms in glibenclamide response, Surendiren and colleagues investigated the influence of these variants on glibenclamide plasma levels and response in 80 SI T2DM patients.

In the study cohort, the genetic polymorphisms of *CYP2C9* and *MDR1* did not influence glibenclamide plasma levels. However, they had an influence on diabetic status with glibenclamide. The variant genotypes of *CYP2C9 *1/*3* and **1/*2* ($p<0.001$) and the normal genotype of *MDR1 CC* ($p<0.01$) were associated with controlled diabetic status in response to therapy, while the normal genotype of *CYP2C9 *1/*1* and the homozygous variant genotype of *MDR1 TT* were associated with uncontrolled diabetic status in therapy response. The study further revealed that the hypoglycemic adverse effects of glibenclamide are associated with its plasma concentration in the presence of normal genotypes for *MDR1*. The genotype of *CYP2C9 *1/*3* and *MDR1* variants did not influence the hypoglycemic adverse effects [60,61].

46.4.3.3 Pioglitazone

Pioglitazone, a peroxisome proliferator–activated receptor gamma (PPARγ) activator, is metabolized by CYP2C8. Pioglitazone response is greatly variable in T2DM patients carrying the genetic variants of the genes *CYP2C9* and *PPARγ*. In view of this interindividual variability, Mathur and coworkers studied pioglitazone monotherapy response in relation to the PPARγ *Pro12Ala* polymorphism in 30 newly diagnosed Indian patients. Since these patients were monomorphic for the *Pro12* allele, association between pioglitazone primary failure rate and the *Pro12Ala* polymorphism was not established [62].

46.4.4 Anticonvulsants

Pharmacoresistance to antiepileptic drugs (phenytoin, carbamazepine, phenobarbital, and valproate) is a major problem in the treatment of epilepsy, likely to be triggered by the combination of genetic and environmental variants. The first Indian study by Soundararajan and colleagues didnot find any association between the *CYP2D6 100C>T* and *2850C>T* polymorphisms and phenytoin-induced toxicity [63]. A later case report revealed phenytoin (DPH) toxicity in a 22-year-old Tamil woman carrying homozygosity for *CYP2C9*3*, a polymorphism associated with a marked decrease in enzyme activity. The woman's condition was reversed on switching to sodium valproate [64]. Studies from our lab indicated a 15.3-fold greater risk of phenytoin toxicity in subjects carrying the *CYP2C9*1/*3* genotype.

In our study, haplotype analysis of *CYP2C9* and *CYP2C19* gene polymorphisms revealed the association of the C-C-G-G haplotype with toxicity [65]. In addition, the contribution of under nourishment and *CYP2C9* genetic variants to phenytoin toxicity were found to be 37.3-fold. This result suggests the additive effect of undernourishment and *CYP2C9* variants in elevating free phenytoin levels individually as well as in combination [66]. Similarly, in an NI population the *CYP2C9*2/*3* polymorphisms were significantly associated with phenytoin toxicity in 259 patients [67].

Three polymorphisms of *MDR1* and *rs3773364 A>G* in synapsin-2 genes were analyzed in NI epileptic patients and neither was found to be associated with multidrug resistance [68,69]. The researchers in this study showed, in two additional studies, the association of the *"A"* allele of 9*SCN2A 56G>A* and *CYP2C9* variants with multidrug resistance [70,71]. They also found significant association of the *GABRA1 IVS11+15 A>G* variant with multidrug resistant but this was not true for the *GABRG2 588C>T*, *GABRA6 1512 T>C*, *GABRB2 1412 C>T*, and *GABRR2 IVS2C>G* genes of GABA (A) receptors and *SCN1AIVS5-91G>A* [72–74].

Balan and colleagues also found no association between *MVP* genes and multidrug resistance [75]. However, a recent report on the pharmacokinetics of phenytoin did find significant association between phenytoin toxicity and the *NAT2* polymorphisms **5A (481C>T)* and **5C (803A>G)*. The pharmacokinetic parameters for phenytoin toxicity in poor metabolizers showed a steady-state concentration of phenytoin compared to intermediate metabolizers and extensive metabolizers [76].

46.4.5 Immunosuppressants

46.4.5.1 Cyclosporine and Tacrolimus

Cyclosporine (CsA) and tacrolimus (Tac) are calcineurin inhibitors; they are the most common immunosuppressive drugs for prevention of graft rejection in organ transplant recipients. CsA and Tac have a narrow therapeutic range; individual dose assessment is thus accomplished by routine therapeutic drug monitoring (TDM). The pharmacodynamics and pharmacokinetics of these agents are influenced mainly by the genetic polymorphisms of *CYP3A4*, *CYP3A5*, and *MDR1*, as Csa and Tac are the substrates of the proteins encoded by these genes [77]. Using existing Indian pharmacogenetic data, researchers examined the association of *CYP3A4*, *CYP3A5*, *MDR1*, and *IL-2* genetic variants with dose requirements of calcineurin inhibitors only in NI renal transplant recipients. In 224 patients on CsA and in 73 on Tac, subjects expressing *CYP3A5* were significantly associated with lower dose-adjusted CsA/Tac blood concentrations; moreover, Tac patients were found to be at a greater risk for allograft rejection. However, no effect of *CYP3A4*1B* on CsA/Tac concentrations was observed [78]. Furthermore, the same researchers in another study reported lower dose-adjusted levels (1, 3, and 6 months) of CsA and Tac and a higher risk of allograft rejection with the wild-type homozygous genotype at *MDR1 2677G>T (GG)* and *3435C>T (CC)* loci, but no correlation was found with *MDR1* haplotypes [79].

An investigation of 100 patients (56 on CsA and 44 on Tac) showed a higher Tac level/dose ratio in recipients with the homozygous mutants *CYP3A5*3/*3* and *MDR1TT* [80]. A prospective study that investigated the impact of *CYP3A5*3 (6986A>G)* on Tac concentrations in 79 renal transplant patients on ketoconazole found significant lower dose-adjusted trough levels in patients carrying at least one **1* allele of *CYP3A5*. In addition, compared to baseline, the dose normalization of Tac was increased by about 112% and 79% in those with and without *CYP3A5*3*, respectively, in response to the metabolic inhibition by ketoconazole [81]. The pharmacogenetics of immunosuppression is not consistent, even though studies have shown strong associations between pharmacogenetic variants and CsA/Tac trough concentrations. The PGx of *CYP3A5* and *MDR1* genes is clinically relevant in optimization and individualized CsA and Tac dosage adjustments. Neverthless, haplotype analyses, along with confounding factors, are likely to be useful in better understanding the variability in the pharmacokinetics of these agents.

46.4.6 Analgesics

46.4.6.1 Opioids

Pain, either acute or chronic, is therapeutically managed by opioid analgesics. Morphine is one of the important and most commonly used opioids in clinical practice. However, variations in genes involved in opioid pharmacokinetics and pharmacodynamics cause interindividual differences in analgesic activity. The only Indian population study on the pharmacogenetic assessment of opioids was carried out with 83 SI women who had undergone a total abdominal hysterectomy. We observed a higher required per-kg dose (0.657±0.107) of morphine in patients with the heterozygous genotype *(CT)* of *MDR1* than was required in those with other genotypes to maintain the same level of analgesia. Similarly, patients with the homozygous variant *(TT)* genotype are at increased risk of developing morphine-induced side effects, such as nausea and vomiting, than are those with the *CC* and *CT* genotypes [82].

46.4.7 Antifolates

46.4.7.1 Methotrexate

Methotrexate (MTX) is a folate inhibitor used as a first-line agent in the treatment of autoimmune diseases and malignancies. As with any drug, a large proportion of patients respond differently to this therapy and develop unpredictable side effects. Such variation may be related to pharmacogenetic variants in the enzymes involved in the folate pathway. Indian studies have examined the pharmacogenetic aspects of methotrexate in rheumatoid arthritis (RA) patients. The first such investigation in 150 NI RA patients produced a negative association between the *MTHFR 677C>T* polymorphism and the therapeutic efficacy of MTX [83]. A second study analyzed two *MTHFR* polymorphisms, *677C>T* and *1298A>C*, to evaluate the relationship but did not find any association with MTX efficacy and toxicity [84]. It has been found that the *CT* heterozygous condition of *MDR1 3435C>T* conferred double the risk of poor MTX response [85]. RA patients with the *FPGS rs1544105 AA* and *AG* and *TYMS rs2853539 AA* genotypes were shown to be poor responders [86].

46.5 CONCLUSIONS AND FUTURE DIRECTIONS

The Council of Scientific and Industrial Research initiative known as the Indian Genome Variation Consortium(IGVC) is the first comprehensive genetic study of Indian populations that addressed genetic diversity, migration, and population relativity by exploring variations in thousands of genes. Although the database named IGV Browser currently harbors allele and genotype frequencies for 4229 SNPs from more than 900 genes in distinct Indian populations, it does not have information regarding ADME pharmagenes and has data only for disease susceptibility markers [87]. Considering this fact—along with the stringent mating practice (endogamy) and rich genetic diversity of Indian populations and the nonavailability of functional studies for these polymorphisms in Indians—there is an immediate need for a functional and molecular study that will characterize the clinically important ADME gene polymorphisms. These characterizations will eventually pave the way for an Indian database with the establishment of a nationwide PGx network among Indian researchers. Such a network will lead to the development of a population-specific PGx chip for Indian populations that should undergo multicenter validation of its clinical benefit and cost effectiveness.

For the precise estimation of allele frequencies, random sampling is essential in population genetics. Further, study subjects should be clustered based on their caste rather than on state of origin; this is imperative for proper population stratification. Moreover, it is essential to generate population-specific normative frequency data before initiating pharmacogenetic association studies. Even more important will be power and sample size calculations based on the normative frequency of the study variants, to be performed before genotype–phenotype correlation studies to avoid spurious false-positive and false-negative results.

Advancements in genetics have given rise to new methods and novel approaches for deciphering differences in treatment outcomes in genomic medicine. SNPs can define only a part of this interindividual variability; pharmacoepigenetics will add another level of understanding in this area. For this reason, studies should focus on epigenetic mechanisms (DNA methylation, miRNA, histone acetylation), copy number variation (CNV) analysis, and gene expression regulation to provide new insights into the inherited traits that cause the variability in pharmacotherapy. The result should be a significant platform for implementing personalized medicine in Indian populations.

REFERENCES

[1] Marshall A. Getting the right drug into the right patient. Nat Biotechnol 1997;15:1249–52.
[2] Hall J, Dennler P, Haller S, Pratsinis A, Sauberli K, Towbin H, et al. Genomics drugs in clinical trials. Nat Rev 2010;9(12):988.
[3] Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. N Engl J Med 2011;364(12):1144–53.
[4] Issa AM. Personalized medicine and the practice of medicine in the 21st century. Mcgill J Med 2007;10:53–7.
[5] Snyder LH. Studies in human inheritance IX. The inheritance of taste deficiency in man. Ohio J Sci 1932;32:436–68.
[6] Vogel F. Moderne problem der human genetik. Ergeb Inn Med U Kinderheilk 1959;12:52–125.
[7] Pharmacogenomic Biomarkers in Drug Labels. <www.fda.gov> [accessed July 2013].
[8] Census of India. <www.censusindia.govt.in/2011census> [accessed June 2013].
[9] Indian Genome Variation Consortium. Genetic landscape of the people of India: a canvas for disease gene exploration. J Genet 2008;87:3–20.
[10] Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Science 2009;461:489–94.
[11] Umamaheswaran G, Krishna Kumar D, Adithan C. Distribution of genetic polymorphisms of genes encoding drug metabolizing enzymes and drug transporters – a review with Indian perspective. Indian J Med Res 2014;139:296–334.
[12] Umamaheswaran G, Dkhar SA, Kalaivani S, Anjana R, Revathy M, Jaharamma M, et al. Haplotype structures and functional polymorphic variants of the drug target enzyme aromatase (CYP19A1) in South Indian population. Med Oncol 2013;30(3):665.
[13] Kumar DK, Shewade DG, Surendiran A, Adithan C. Genetic variation and haplotype structure of the gene vitamin K epoxide reductase complex, subunit 1 in the Tamilian population. J Pharmacol Pharmacother 2013;4(1):53–8.
[14] Minhas S, Setia N, Pandita S, Saxena R, Verma IC, Aggarwal S. Prevalence of CYP2C8 polymorphisms in a North Indian population. Genet Mol Res 2013;12(3):2260–6.
[15] Umamaheswaran G, Krishna Kumar D, Kayathiri D, Rajan S, Shewade DG, Dkhar SA, et al. Inter and intra-ethnic differences in the distribution of the molecular variants of TPMT, UGT1A1 and MDR1 genes in the South Indian population. Mol Biol Rep 2012;39(5):6343–51.
[16] Nahar R, Deb R, Saxena R, Puri RD, Verma IC. Variability in CYP2C9 allele frequency: a pilot study of its predicted impact on warfarin response among healthy South and North Indians. Pharmacol Rep 2013;65(1):187–94.
[17] International HapMap project, <http://hapmap.ncbi.nlm.nih.gov/> [accessed July 2013].
[18] Dayakar S, Goud KI, Reddy TP, Rao SP, Sesikeran SB, Sadhnani M. Sequence variation of the methylene tetrahydrofolate reductase gene (677C>T and 1298A>C) and traditional risk factors in a South Indian population. Genet Test Mol Biomarkers 2011;15(11):765–9.
[19] Saraswathy KN, Asghar M, Samtani R, Murry B, Mondal PR, Ghosh PK, et al. Spectrum of MTHFR gene SNPs C677T and A1298C: a study among 23 population groups of India. Mol Biol Rep 2012;39(4):5025–31.
[20] Asghar M, Kabita S, Kalla L, Murry B, Saraswathy KN. Prevalence of MTHFR, factor V, ACE and APOE gene polymorphisms among Muslims of Manipur India. Ann Hum Biol 2013;40(1):83–7.
[21] Ramu P, Umamaheswaran G, Shewade DG, Swaminathan RP, Dutta TK, Balachander J, et al. Candidate gene polymorphisms of renin angiotensin system and essential hypertension in a South Indian Tamilian population. Int J Hum Genet 2011;11(1):31–40.
[22] Tripathi G, Dharmani P, Khan F, Sharma RK, Pandirikkal V, Agrawal S. High prevalence of ACE DD genotype among north Indian end stage renal disease patients. BMC Nephrol 2006;7:15.
[23] Zhang YM, Zhang LY, Wang KQ, Ge JB. Distribution of angiotensin converting enzyme gene polymorphism among Northern Hans, Dahurs, and Ewenks. Acta Pharmacol Sin 2001;22(8):747–50.
[24] Dhandapany PS, Sadayappan S, Xue Y, Powell GT, Rani DS, Nallar P, et al. A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia. Nat Genet 2009;41:187–91.
[25] Rani DS, Carlus SJ, Poongothai J, Jyothi A, Pavani K, Gupta NJ, et al. CAG repeat variation in the mtDNA polymerase gamma is not associated with oligoasthenozoospermia. Int J Androl 2009;32:647–55.
[26] Mehrotra S, Oommen J, Mishra A, Sudharshan M, Tiwary P, Jamieson SE, et al. No evidence for association between SLC11A1 and visceral leishmaniasis in India. BMC Med Genet 2011;12:71.
[27] Agarwal SS, Phadke SR, Phadke RV, Das SK, Singh GK, Sharma JP, et al. Handigodu disease a radiological study. A new variety of spondyloepi(meta)physeal dysplasia of the autosomal dominant type. Skeletal Radiol 1994;23:611–9.
[28] Nalini A, Yamini BK, Gayatri N, Thennarasu K, Gope R. Familial Madras motor neuron disease (FMMND) study of 15 families from southern India. J Neurol Sci 2006;250:140–6.
[29] Pandit JJ, Gopa S, Arora J. A hypothesis to explain the high prevalence of pseudo-cholinesterase deficiency in specific population groups. Eur J Anaesthesiol 2011;28:550–2.
[30] Hertz DL, McLeod HL, Irvin Jr WJ. Tamoxifen and CYP2D6: a contradiction of data. Oncologist 2012;17(5):620–30.

[31] Damodaran SE, Pradhan SC, Umamaheswaran G, Kadambari D, Reddy KS, Adithan C. Genetic polymorphisms of CYP2D6 increase the risk for recurrence of breast cancer in patients receiving tamoxifen as an adjuvant therapy. Cancer Chemother Pharmacol 2012;70(1):75–81.

[32] Damodaran SE, Pradhan SC, Umamaheswaran G, Kadambari D, Reddy KS, Adithan C. No association between *CYP2C19* variant alleles and recurrence of breast cancer in Tamilian patients on adjuvant tamoxifen therapy. J Cancer Res Therap 2012;8(1):S30.

[33] Kalyan Kumar Ch, Mohan Reddy N, Laxmi A, Adithya V, Tabassun SN, Pallavi Y, et al. Clinical impact of CYP2D6 and SULT1A1 polymorphisms and tamoxifen with breast cancer. Open Access Sci Rep 2013;2(3):683.

[34] Hadfield KD, Newman WG. Pharmacogenetics of aromatase inhibitors. Pharmacogenomics 2012;13(6):699–707.

[35] George J, Dharanipragada K, Krishnamachari S, Chandrasekaran A, Sam SS, Sunder E. A single-nucleotide polymorphism in the MDR1 gene as a predictor of response to neoadjuvant chemotherapy in breast cancer. Clin Breast Cancer 2009;9(3):161–5.

[36] Vishnukumar S, Umamaheswaran G, Anichavezhi D, Indumathy S, Adithan C, Srinivasan K, et al. P-glycoprotein expression as a predictor of response to neoadjuvant chemotherapy in breast cancer. Indian J Cancer 2013;50(3):195–9.

[37] Mishra A, Chandra R, Mehrotra PK, Bajpai P, Agrawal D. Glutathione S-transferase M1 and T1 polymorphism and response to neoadjuvant chemotherapy (CAF) in breast cancer patients. Surg Today 2011;41(4):471–6.

[38] Chaturvedi P, Tulsyan S, Agarwal G, Lal P, Agarwal S, Mittal RD, et al. Influence of ABCB1 genetic variants in breast cancer treatment outcomes. Cancer Epidemiol 2013;37(5):754–61.

[39] Tulsyan S, Chaturvedi P, Agarwal G, Lal P, Agrawal S, Mittal RD, et al. Pharmacogenetic influence of GST polymorphisms on anthracycline-based chemotherapy responses and toxicity in breast cancer patients: a multi-analytical approach. Mol Diagn Ther 2013;17(6):371–9.

[40] Dorababu P, Nagesh N, Linga VG, Gundeti S, Kutala VK, Reddanna P, et al. Epistatic interactions between thiopurine methyltransferase (TPMT) and inosine triphosphate pyrophosphatase (ITPA) variations determine 6-mercaptopurine toxicity in Indian children with acute lymphoblastic leukemia. Eur J Clin Pharmacol 2012;68(4):379–87.

[41] Dorababu P, Naushad SM, Linga VG, Gundeti S, Nagesh N, Kutala VK, et al. Genetic variants of thiopurine and folate metabolic pathways determine 6-MP-mediated hematological toxicity in childhood ALL. Pharmacogenomics 2012;13(9):1001–8.

[42] Kapoor G, Sinha R, Naithani R, Chandgothia M. Thiopurine S-methyltransferase gene polymorphism and 6-mercaptopurine dose intensity in Indian children with acute lymphoblastic leukemia. Leuk Res 2010;34(8):1023–6.

[43] Subraja K, Dkhar SA, Priyadharsini R, Ravindra BK, Shewade DG, Satheesh S, et al. Genetic polymorphisms of CYP2C19 influences the response to clopidogrel in ischemic heart disease patients in the South Indian Tamilian population. Eur J Clin Pharmacol 2013;69(3):415–22.

[44] Subraja K. Effect of CYP2C19 and MDR1 genetic polymorphisms on antiplatelet effects of clopidogrel in clinical settings; 2012 Dissertation.

[45] Priyadharsini R. Effect of CYP3A5 and P2Y12 genetic polymorphisms on platelet aggregation in patients on treatment with clopidogrel; 2012 Dissertation.

[46] Kar R, Meena A, Yadav BK, Yadav R, Kar SS, Saxena R. Clopidogrel resistance in North Indian patients of coronary artery disease and lack of its association with platelet ADP receptors P2Y1 and P2Y12 gene polymorphisms. Platelets 2013;24(4):297–302.

[47] Shalia KK, Shah VK, Pawar P, Divekar SS, Payannavar S. Polymorphisms of MDR1, CYP2C19 and P2Y12 genes in Indian population: effects on clopidogrel response. Indian Heart J 2013;65(2):158–67.

[48] Pavani A, Naushad SM, Rupasree Y, Kumar TR, Malempati AR, Pinjala RK, et al. Optimization of warfarin dose by population-specific pharmacogenomic algorithm. Pharmacogenomics J 2012;12(4):306–11.

[49] Pavani A, Naushad SM, Mishra RC, Malempati AR, Pinjala R, Kumar TR, et al. Retrospective evidence for clinical validity of expanded genetic model in warfarin dose optimization in a South Indian population. Pharmacogenomics 2012;13(8):869–78.

[50] Krishna Kumar D, Shewade DG, Loriot MA, Beaune P, Balachander J, Sai Chandran BV, et al. Effect of CYP2C9, VKORC1, CYP4F2 and GGCX genetic variants on warfarin maintenance dose and explicating a new pharmacogenetic algorithm in South Indian population. Eur J Clin Pharmacol 2014;70(1):47–56.

[51] Gaikwad T, Ghosh K, Kulkarni B, Kulkarni V, Ross C, Shetty S. Influence of CYP2C9 and VKORC1 gene polymorphisms on warfarin dosage, over anticoagulation and other adverse outcomes in Indian population. Eur J Pharmacol 2013;710(1–3):80–4.

[52] Shalia KK, Doshi SM, Parikh S, Pawar PP, Divekar SS, Varma SP, et al. Prevalence of VKORC1 and CYP2C9 gene polymorphisms in Indian population and its effect on warfarin response. J Assoc Physicians India 2012;60:34–8.

[53] Rathore SS, Agarwal SK, Pande S, Singh SK, Mittal T, Mittal B. Therapeutic dosing of acenocoumarol: proposal of a population specific pharmacogenetic dosing algorithm and its validation in North Indians. PLoS One 2012;7(5):e37844.

[54] Kaur A, Khan F, Agrawal SS, Kapoor A, Agarwal SK, Phadke SR. Cytochrome P450 (CYP2C9*2,*3) and vitamin-K epoxide reductase complex (VKORC1 -1639G<A) gene polymorphisms and their effect on acenocoumarol dose in patients with mechanical heart valve replacement. Indian J Med Res 2013;137(1):203–9.

[55] Nahar R, Saxena R, Deb R, Verma IC. Pharmacogenetic typing for oral anti-coagulant response among factor V Leiden mutation carriers. Indian J Hum Genet 2012;18(3):326–31.

[56] Krishna Kumar D, Madhan S, Balachander J, Sai Chandran BV, Thamijarassy B, Adithan C. Effect of CYP2C9 and VKORC1 genetic polymorphisms on mean daily maintenance dose of acenocoumarol in South Indian patients. Thromb Res 2013;131(4):363–7.

[57] Poduri A, Khullar M, Bahl A, Sehrawat BS, Sharma Y, Talwar KK. Common variants of HMGCR, CETP, APOAI, ABCB1, CYP3A4, and CYP7A1 genes as predictors of lipid-lowering response to atorvastatin therapy. DNA Cell Biol 2010;29(10):629–37.

[58] Srivastava A, Garg N, Srivastava A, Srivastava K, Mittal B. Effect of genetic variant (rs11887534) in ABCG8 gene in coronary artery disease and response to atorvastatin therapy. Dis Markers 2010;28(5):307–13.

[59] Saranya V. Influence of ATM and SLC22A1 gene polymorphisms on metformin response in type 2 diabetes mellitus patients of South India; 2014 Dissertation.

[60] Surendiran A, Pradhan SC, Agrawal A, Subrahmanyam DK, Rajan S, Anichavezhi D, et al. Influence of CYP2C9 gene polymorphisms on response to glibenclamide in type 2 diabetes mellitus patients. Eur J Clin Pharmacol 2011;67(8):797–801.

[61] Surendiran. Effect of the CYP2C9 and MDR1 polymorphisms on plasma levels of glibenclamide in patients with non-insulin dependent diabetes mellitus; 2011 Dissertation.

[62] Mathur SK, Rathore R, Chandra S, Sharma BB, Sharma L, Mathur P, et al. Primary pioglitazone failure in Asian Indian diabetics is not related to common Pro12Ala polymorph of PPAR-gamma gene. Indian J Physiol Pharmacol 2009;53(2):175–80.

[63] Soundararajan CC, Adhin A, Tripathi M, Prasad HK, Jain S, Vivekanandhan S. Association of CYP2D6 100 C>T and 2850 C>T polymorphisms with generalized tonic clonic seizures among Indians. J Neural Transm 2006;113(12):1873–6.

[64] Ramasamy K, Narayan SK, Chanolean S, Chandrasekaran A. Severe phenytoin toxicity in a CYP2C9*3*3 homozygous mutant from India. Neurol India 2007;55(4):408–9.

[65] Kesavan R, Narayan SK, Adithan C. Influence of CYP2C9 and CYP2C19 genetic polymorphisms on phenytoin-induced neurological toxicity in Indian epileptic patients. Eur J Clin Pharmacol 2010;66(7):689–96.

[66] Ramasamy K, Narayan SK, Shewade DG, Chandrasekaran A. Influence of CYP2C9 genetic polymorphism and undernourishment on plasma-free phenytoin concentrations in epileptic patients. Ther Drug Monit 2010;32(6):762–6.

[67] Thakkar AN, Bendkhale SR, Taur SR, Gogtay NJ, Thatte UM. Association of CYP2C9 polymorphisms with phenytoin toxicity in Indian patients. Neurol India 2012;60(6):577–80.

[68] Lakhan R, Misra UK, Kalita J, Pradhan S, Gogtay NJ, Singh MK, et al. No association of ABCB1 polymorphisms with drug-refractory epilepsy in a north Indian population. Epilepsy Behav 2009;14(1):78–82.

[69] Lakhan R, Kalita J, Misra UK, Kumari R, Mittal B. Association of intronic polymorphism rs3773364 A>G in synapsin-2 gene with idiopathic epilepsy. Synapse 2010;64(5):403–8.

[70] Lakhan R, Kumari R, Misra UK, Kalita J, Pradhan S, Mittal B. Differential role of sodium channels SCN1A and SCN2A gene polymorphisms with epilepsy and multiple drug resistance in the north Indian population. Br J Clin Pharmacol 2009;68(2):214–20.

[71] Lakhan R, Kumari R, Singh K, Kalita J, Misra UK, Mittal B. Possible role of CYP2C9 and CYP2C19 single nucleotide polymorphisms in drug refractory epilepsy. Indian J Med Res 2011;134:295–301.

[72] Kumari R, Lakhan R, Kalita J, Misra UK, Mittal B. Association of alpha subunit of GABAA receptor subtype gene polymorphisms with epilepsy susceptibility and drug resistance in north Indian population. Seizure 2010;19(4):237–41.

[73] Kumari R, Lakhan R, Kalita J, Garg RK, Misra UK, Mittal B. Potential role of GABAA receptor subunit; GABRA6, GABRB2 and GABRR2 gene polymorphisms in epilepsy susceptibility and pharmacotherapy in North Indian population. Clin Chim Acta 2011;412(13–14):1244–8.

[74] Kumari R, Lakhan R, Kumar S, Garg RK, Misra UK, Kalita J, et al. SCN1AIVS5-91G>A polymorphism is associated with susceptibility to epilepsy but not with drug responsiveness. Biochimie 2013;95(6):1350–3.

[75] Balan S, Lekshmi S, Radha K, Sathyan S, Vijai J, Banerjee M, et al. Major vault protein (MVP) gene polymorphisms and drug resistance in mesial temporal lobe epilepsy with hippocampal sclerosis. Gene 2013;526(2):449–53.

[76] Murali M, Manjari T, Madhuri B, Raghavan S, Jain DC, Vivekanandhan S. Genetic polymorphism of NAT2 metabolizing enzymes on phenytoin pharmacokinetics in Indian epileptic patients developing toxicity. CNS Neurosci Ther 2012;18(4):350–8.

[77] Staatz CE, Goodman LK, Tett SE. Effect of CYP3A and ABCB1 single nucleotide polymorphisms on the pharmacokinetics and pharmacodynamics of calcineurin inhibitors: Part II. Clin Pharmacokinet 2010;49(4):207–21.

[78] Singh R, Srivastava A, Kapoor R, Sharma RK, Mittal RD. Impact of CYP3A5 and CYP3A4 gene polymorphisms on dose requirement of calcineurin inhibitors, cyclosporine and tacrolimus, in renal allograft recipients of North India. Naunyn Schmiedebergs Arch Pharmacol 2009;380(2):169–77.

[79] Singh R, Srivastava A, Kapoor R, Mittal RD. Do drug transporter (ABCB1) SNPs influence cyclosporine and tacrolimus dose requirements and renal allograft outcome in the posttransplantation period? J Clin Pharmacol 2011;51(4):603–15.

[80] Ashavaid T, Raje H, Shalia K, Shah B. Effect of gene polymorphisms on the levels of calcineurin inhibitors in Indian renal transplant recipients. Indian J Nephrol 2010;20(3):146–51.

[81] Chandel N, Aggarwal PK, Minz M, Sakhuja V, Kohli KK, Jha V. CYP3A5*1/*3 genotype influences the blood concentration of tacrolimus in response to metabolic inhibition by ketoconazole. Pharmacogenet Genomics 2009;19(6):458–63.

[82] Shanil J. Association of *MDR1* gene polymorphism with morphine requirement for patient controlled analgesia after total abdominal hysterectomy in a south Indian population; 2010 Dissertation.

[83] Aggarwal P, Naik S, Mishra KP, Aggarwal A, Misra R. Correlation between methotrexate efficacy and toxicity with C677T polymorphism of the methylenetetrahydrofolate gene in rheumatoid arthritis patients on folate supplementation. Indian J Med Res 2006;124(5):521–6.

[84] Ghodke Y, Chopra A, Joshi K, Patwardhan B. Are thymidylate synthase and methylene tetrahydrofolate reductase genes linked with methotrexate response (efficacy, toxicity) in Indian (Asian) rheumatoid arthritis patients? Clin Rheumatol 2008;27(6):787–9.

[85] Sharma S, Das M, Kumar A, Marwaha V, Shankar S, Aneja R, et al. Interaction of genes from influx-metabolism-efflux pathway and their influence on methotrexate efficacy in rheumatoid arthritis patients among Indians. Pharmacogenet Genomics 2008;18(12):1041–9.

[86] Sharma S, Das M, Kumar A, Marwaha V, Shankar S, Singh P, et al. Purine biosynthetic pathway genes and methotrexate response in rheumatoid arthritis patients among north Indians. Pharmacogenet Genomics 2009;19(10):823–8.

[87] IGVdb portal. <www.igvdb.res.in/> [accessed August 2013].

This page is intentionally left blank

Glossary

Agonist stimulates or activates the receptor to produce an effect

Allele one version of a gene at a given locus along a chromosome

Allosteric modulation indirect modulation of the effects of an agonist by binding to a separate binding site, leading to an increase or decrease the receptor's affinity for agonist

Alpha helix and beta sheet common secondary protein structures

Alternative splicing a regulatory mechanism by which multiple mRNA transcripts are generated from a single gene

Amino acid substitution matrix a matrix describing how common each type of amino acid change is, to facilitate scoring of similarity between protein sequences

Andersen-Tawil syndrome a condition characterized by episodic paralysis, ventricular arrhythmias, and prolonged QT interval. Approximately 60% of individuals with Andersen-Tawil have a mutation in KCNJ2. This encodes the inward rectifier potassium channel 2 protein, Kir2.1

Antagonist prevents the action of an agonist but does not have any effect itself

Argonaute 2 a protein that binds to small RNAs and is a component of RISC. In mammals, there are four Ago proteins, Ago1–4, which function in translational repression; only Ago2 has slicer activity

Band During cell division, the specific staining pattern (light and dark stripes) is observed for chromosomes. Such patterns can help in identifying chromosomes and evaluating their structure

Biallelic pertaining to both alleles of a gene (in a heterozygote)

Binary fission a type of cell division that applies to prokaryotes, such as bacteria. One cell is divided in two, with two complete and identical genomes present in the two daughter cells

Biomarker a measurable characteristic that is an indicator of normal biologic processes, pathogenic processes, and/or response to therapeutic or other interventions

BLAST a popular algorithm for sequence alignment

Bootstrapping the scoring of a tree or nodes of a tree by resampling and counting outcomes of each pseudoreplicate dataset

Cascade screening method of ascertaining gene carriers by testing the patient's close family members

Centromere a constricted chromosomal region that holds sister chromatids together during cell division. The centromere also separates a chromosome into a short arm and a long arm

Character a heritable feature that varies among individuals

Chromatid during the S phase of the cell cycle, each chromosome is duplicated. Following replication, each chromosome consists of two identical sister chromatids, which contain matched copies of the DNA sequence at every locus

Chromatin fiber the 30 nm coil of DNA and proteins that forms the basic structure of chromatin

Clinical annotation a genotype-based, evidence-rated summary of PharmGKB's literature evidence for a particular variant and drug phenotype association

Companion diagnostics an *in vitro* diagnostic device that provides information essential for the safe and effective use of a corresponding therapeutic product

Copy number variant form of DNA variation in which a certain sequence is present in differing numbers of copies, and is different across individuals. Many CNVs are nonpathogenic common variants

Copy number variants deletions, duplications, and insertions in the germline genome that are >1 kb in length

Cytochrome P450 a large superfamily of heme-thiolate proteins involved in the metabolism of a wide variety of exogenous and endogenous compounds

Cytochrome P450s the human cytochrome P450 superfamily, which comprises 57 genes. P450 sequences that display greater than 40% amino acid identity are classified in the same family; sequences that have greater than 55% identity are classified in the same subfamily. Members of CYP1–3 subfamilies are involved in the metabolism of xenobiotics, including drugs

Cytokinesis the division of the cytoplasm. Cytokinesis immediately follows the division of the nucleus

Cytoskeleton protein filaments within cells that maintain cell shape and cell movement

Delimiter separated value a plain text file in which rows are separated by EOL characters and columns by a specified delimiter (e.g., tab or comma)

De novo structure prediction the prediction of a protein's structure without the use of templates

Desensitization repeated exposure to a drug leading to a diminishing effect

Dicer an endoribonuclease that cleaves pre-miRNAs into 20–25-nucleotide, double-stranded miRNAs

Digenic Pertaining to two genes interacting to produce the phenotype. Often the mode of inheritance is Mendelian with another gene modulating disease severity

Discrete method a method that considers differences in nucleotides directly

Distance method a method that considers only the number of differences between nodes

DNA the molecule that encodes the genes responsible for the structure and function of an organism. Genetic information carried by DNA molecules is transmitted from one generation to the next

DNA methylation methyl groups, consisting of one carbon atom and three hydrogen atoms, attach to DNA molecules at cytosine bases, resulting in reduced transcription of the gene

Dot-matrix a graph representing similarities between two sequences

Endomembrane system all membranes, which are suspended in the cytoplasm within a cell, can divide the cell into structural and functional compartments/organelles. In a eukaryotic cell, endomembrane organelles include the nuclear envelope, the

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00056-6

© 2014 Elsevier Inc. All rights reserved.

endoplasmic reticulum, the Golgi apparatus, lysosomes, vacuoles, vesicles, and the cell membrane. The endomembrane system does not contain mitochondrial membranes

Epigenomics the study of global gene expression changes caused by epigenetic processes that do not involve any change in the underlying DNA sequence

Epistasis the expression of one gene depends on the expression of one or more independently inherited genes

Ethnic fractionalization (EF) the probability that two individuals selected at random from a country will be from different ethnic groups

Eukaryote an organism whose cells contain a nucleus and other membrane-bound components. Many unicellular organisms, such as protozoa, are eukaryotes. All multicellular organisms—animals, plants, and fungi—are eukaryotes

Exome the part of the genome that corresponds to exons, which are the coding sequences of DNA present in mature messenger RNA. The exome constitutes about 1% of the total human genome

Expression quantitative trait loci genetic loci that regulate gene expression

F-statistics (F_{ST}) the proportion of the total genetic variance contained in a subpopulation (subscript S) relative to the total genetic variance (subscript T). Values can range from 0 to 1. High F_{ST} implies a considerable degree of differentiation among populations. F_{IS} (inbreeding coefficient) is the proportion of the variance in the subpopulation contained in an individual. High F_{IS} implies a high degree of inbreeding

Gamete a cell/gamete that carries half the genetic information of an individual fuses with another cell/gamete in organisms that reproduce sexually. There are two types of gametes: a sperm is produced by a male and an ovum is produced by a female

Gap penalty a penalty on a sequence alignment score, made for the introduction of a gap

Gene the basic physical unit of heredity. For example, a segment of DNA that is arranged in a linear manner along a chromosome is a gene. A gene codes for a specific protein that has a particular characteristic or function

Gene ontology a standardized way of classifying genes and their products

Genome the entire genetic material found in a cell. In the case of humans, the genome includes 23 pairs of chromosomes in the nucleus and the chromosomes contained in a cell's mitochondria

Genome-wide association study (GWAS) analysis of DNA markers across the genome to identify variants associated with particular phenotypes

Genotype the genetic make-up of a cell or organism. Also, a set of alleles inherited at a locus

Haplotype a set of alleles or single-nucleotide polymorphisms found on the same chromosome, tending to be inherited together

HAT (sleeping sickness) Human African trypanosomiasis is the correct term for human infections caused by the African trypanosomes of the *T. brucei* species: *Trypanosoma brucei gambiense* and *T. b. rhodesiense*. The disease is commonly referred to as sleeping sickness when in its second, CNS involved stage. At this stage, trypanosomes cause various perturbations to the neurological system, including the changes to sleep/wake patterns that led to the disease acquiring its common name

Heuristic a "rule-of-thumb" method; not as accurate as, but faster than, an exact method

Homology relation by descent, inferred by similarity. Often used synonymously with 'similarity'

Homology modeling the use of similar protein sequences with known structures in protein modeling

Imputation in statistics, imputation is a process of replacing missing data with other estimated probable values. In genetics, imputation is a method used to fill in missing genotypes in a study dataset

Inverse agonist produces an opposite effect to that of the agonist

Kinetochore a protein structure on chromatids, allowing the spindle fibers to attach and pull the sister chromatids apart during cell division

Label a description of a drug product, including what it is used for, who should take it, side effects, instructions for use, and safety information

Linkage disequilibrium nonrandom association of alleles at two or more loci

Linked genes genes on the same chromosome tend to be segregated and inherited together

Locus the location or the physical site of an individual gene or DNA sequence on a chromosome

Minor allele frequency the frequency of the least common allele in a given population

MIM number an ID from OMIM specifying a disease with a genetic component

miRISC an RNA-induced silencing complex that contains mature miRNA and guides its inhibitory function to the target mRNA

Molecular evolution the study of the evolution of DNA, RNA, or proteins

Monogenic Pertaining to a disorder occurring as a direct consequence of single gene defects

Multiple-sequence alignment an alignment of several sequences, often used in phylogenetics and protein modeling

Mutation an alteration in a DNA sequence from its natural state. The effect of any mutation might be deleterious or benign

Next-generation sequencing a high-throughput sequencing technology relative to capillary electrophoresis–based Sanger sequencing, producing thousands or millions of sequences simultaneously. This technology can be applied to DNA analysis, RNA analysis, and gene regulation analysis

Noncoding RNA a biologically functional RNA actively transcribed from a gene that contains no open reading frame

Organ bath standard for *in vitro* pharmacodynamic dose-response experiments involving suspension of an isolated section of organ/tissue in a controlled solution and sensors to measure the contraction or relaxation of the tissue when drugs are added

Organelle a subcellular structure within a cell. An organelle is usually enclosed by its own membrane for executing its specific function independently of other organelles in a cell, particularly a eukaryotic cell

Organism any living system that has the ability to grow, reproduce, maintain homeostasis within itself, and function independently. An organism may be either unicellular or multicellular. All organisms on Earth can be further divided into eukaryotes and prokaryotes

Ortholog a corresponding gene in different species that has descended from the same ancestral gene

Orthology homology where the sequences have deviated by speciation

Paralog a gene that is derived from the same ancestral gene but reside in different locations in the same genome

Paralogy homology where the sequences have deviated by gene duplication

Partial agonist stimulates the receptor to a limited extent but also prevents any further stimulation by naturally occurring agonists

Pathway an illustration of the pharmacokinetics or pharmacodynamics of a PGx-relevant drug, accompanied by informative text

PDB file a file listing coordinates for the atoms of a protein

Penetrance probability of a characteristic being present given a certain genotype

Pharmacodynamics the study of the physiological and molecular effects of drugs on their recipient and the relationship between plasma drug concentration and drug effect

Pharmacogene a gene involved in drug response

Pharmacogenetics (PGt) the study of variations in DNA sequence as related to drug response

Pharmacogenomics (PGx) the study of variations in DNA and RNA characteristics as related to drug response

Phenotype an organism's observable characteristic or trait, such as biochemical or physiological properties or behavior. Phenotypes result from the expression of an organism's genes as well as the influence of environmental factors and the interactions between the two

Phylogenetics the identification of evolutionary relationships between organisms or between genes

Pleiotropy multiple physical effects caused by a single gene. The underlying mechanism of pleiotropy may be a gene involved in different metabolic pathways that contribute to different phenotypes

φ and ψ angles rotation angles of certain bonds of a protein backbone

Polymorphism a variation in DNA sequence that is present at an allele frequency of $\geq 1\%$ of the population

Prokaryote an organism lacking a membrane-enclosed nucleus. Most prokaryotes are unicellular. No membrane-bound organelles are present in the prokaryotic cells. The intracellular components of a prokaryotic cell are gathered together and enclosed only by the cell membrane

Protein modeling the estimation of a protein's structure using its sequence

Receptor a component of the organism with which a drug interacts

Recombination the new combinations of the traits in offspring that are inherited from its parents, now more commonly known as the exchange of a DNA segment between two homologous chromosomes during meiosis, which causes new combinations of hereditary material in a gamete

Relational database a database with a complex structure, driven by a relational database management system (RDBMS)

RNA a molecule synthesized from the DNA template, with the alternating sugar ribose to replace the deoxyribose present in a DNA molecule. Different types of RNA exist in a cell: messenger RNA, ribosomal RNA, transfer RNA, and some small RNAs

rsID (reference SNP identification) an identification number assigned by the National Center for Biotechnology Information (NCBI) SNP database (dbSNP) to a cluster of SNPs that map to the same location

Second messenger a molecule that relays signals from receptors to the cell surface target

Sequencing the technique for determining the exact DNA sequence or nucleotides (A, C, G, and T). DNA-sequencing technology is developing fast and becoming less expensive. For example, whole-genome and whole-exome sequencing is now available for studying the association between a DNA sequence and a disease

Serum resistance–associated (*SRA*) gene The *SRA* is found in *Trypanosoma brucei rhodesiense*. It is derived from a *VSG* gene and is expressed from a single expression site; hence, trypanosomes carrying the gene are only resistant to serum when the particular *VSG* expression site with which they are associated is expressed. Expression sites are found at the telomeres of the parasite's chromosomes and contain a strong RNA polymerase 1 promoter that allows expression of a single *VSG* gene at a time, in addition to several expression site–associated genes (*ESAGS*) that are also found there

Single nucleotide polymorphism a change in a single nucleotide at a genetic locus with a minor allele frequency of 1% or higher in a population

Somatic cell the cell forms the body of an organism. For example, in humans, somatic cells form all internal organs, skin, bones, and connective tissues. Somatic cells are diploid

Synapsis the pairing of two homologous chromosomes at the early stage of meiosis. This allows the matching of homologous chromosomal pairs and prepares for possible crossover between them

Syncope transient loss of consciousness

Tachyphylaxis decrease in response to a drug following repeated administration

***T. b. gambiense*–specific glycoprotein gene (*TGSGP*)** *TGSGP*, like *SRA* in *T. b. rhodesiense*, is found only in one subspecies. It is also a mutated version of a variant surface glycoprotein gene and can confer resistance against trypanosome lytic factors in human serum, apparently by influencing membrane fluidity in ways that affect uptake of TLF

Therapeutic index the relationship between therapeutic and toxic doses of a drug

Threading aligning a protein sequence to known structures

Time complexity describes the relationship between the input size n and the time taken for an algorithm to run

Timothy syndrome Rare childhood condition that leads to cardiac arrhythmia, cognitive abnormalities, and syndactyly. It is caused by a mutation in the Cav1.2 calcium channel

Torsades de Pointes (TdP) distinctive form of polymorphic ventricular tachycardia characterized by twisting of the QRS complexes around the isoelectric line. It is associated with QT prolongation

Trait a variant of a character, which may be determined by genes, the environment, and the interactions between the two

***trans* configuration** the two alpha carbon atoms of the connected amino acids are on the opposite sides of the peptide bond

Trypanosome lytic factors (TLFs) TLFs are components of human blood that are capable of killing trypanosomes. Initially identified as being part of the high-density lipoprotein fraction of plasma, it was shown that a key individual component was apolipoprotein L1 (APOL1). APOL1 is capable of lysing trypanosomes other than those subspecies that have evolved molecules that neutralize its activity. Other components—such as a haptoglobin-like protein that is part of the complex particle that also contains APOL1—are required for the binding of these particles to the trypanosome and contribute to activity since specific receptors are required for this binding

Tsetse fly The tsetse fly is a dipteran insect of the genus *Glossina*. Various species can transmit trypanosomes, which replicate within the tsetse fly midgut as procyclic forms. Some of these forms migrate to the salivary gland, where they transform to metacyclic forms that are preadapted to survive within the insect host

Variant surface glycoproteins (VSGs) AQ8 VSGs are expressed one at a time, and acquire a glycosylphosphatidyl inositol (GPI) anchor that links them to the plasma membrane. The trypanosome has up to a thousand different *VSG* genes, but expresses only one of them at a time. The proteins form dimers that pack tightly at the parasite's surface and prevent host immunological proteins, including components of the complement complex and immunoglobulins, from reaching the cell surface. Eventually the host produces antibodies that bind to a given antigen type, and cells carrying that VSG are killed. However, a small number of parasites change expression to a new VSG that is not recognized by those antibodies. These cells then proliferate until a specific antibody response is raised to them. In this way, the parasites can evade the immune system *ad infinitum*

Ventricular fibrillation (VF) rapid, chaotic, nonrepetitive wave form in which there is uncoordinated contraction of the cardiac muscle of the ventricles. VF is clinically associated with loss of effective blood circulation and, if not immediately treated, death

Ventricular tachycardia (VT) sequence of three or more beats in a row, with wide QRS complex (QRS > 120 ms) at a ventricular rate exceeding 100 beats/min.

VIP (very important pharmacogene) summary a pharmacogenetic-based overview of an important gene involved in drug response

VNTR the location in a genome where a short nucleotide sequence is organized as a tandem repeat, often showing variations in length between individuals

Yotiao A-kinase (PRKA) anchor protein 9 that directly associates with the slow outward potassium ion current (I_{Ks}) in heart.

Xenobiotic a foreign chemical, which can be manufactured or natural, such as a drug, an industrial chemical, a pollutant, a pyrolysis product in cooked food, an alkaloid, or a secondary plant metabolite.

XML a text file with a hierarchical structure

Acronyms and Abbreviations

2D-PAGE	2D polyacrylamide electrophoresis
3C	chromosome conformation capture
5-FU	5-fluorouracil
6-MP	6-mercaptopurine
A1AT	alpha 1 anti-trypsin
AA	arachidonic acid
AACE	American Association of Clinical Endocrinologists
Aβ	amyloid beta EAD early afterdepolarization
ABCB1	ATP-binding cassette, subfamily B (MDR/TAP), member 1
AC	amoxicillin-clavulanate
ACE	American College of Endocrinology
ACE	angiotensin-converting enzyme
ACS	acute coronary syndromes
AD	antidepressant
AD	Alzheimer's disease
ADA	American Diabetes Association
ADAMTS13	a disintegrin and metalloproteinase with a thrombospondin type 1 motif, member 13
ADE	adverse drug event
ADME	absorption, distribution, metabolism, and excretion
ADP	adenosine diphosphate
ADR	adverse drug reaction
ADRs	adverse drug reactions
AF	amniotic fluid
AFP	alphafetoprotein
AfSHG	African Society for Human Genetics
AG	agamous
AGEP	acute generalized exanthematous pustulosis
Ago2	argonaute 2
AGTR	angiotensin receptor 1
AHR	abacavir hypersensitivity reaction
AhR	aryl hydrocarbon receptor
AHS	abacavir hypersensitivity syndrome
AKAP	A-kinase anchor protein 9
ALK	anaplastic lymphoma kinase
ALL	acute lymphoblastic leukemia
ANK2	ankyrin 2
APC	activated protein C
APD	action potential duration
APEX	absolute protein expression
aQTL	activity quantitative trait loci
ARNT	aryl hydrocarbon receptor nuclear translocator
AS	alternative splicing
AsO	antisense oligonucleotide
ASPE	Allele Specific Primer Extension
ATM	ataxia telangiectasia mutated
ATP	adenosine triphosphate
ATX	Ataxin
AUC	area under the concentration-time curve
AZA	azathioprine
BF	Bayes factor
bFGF	basic fibroblast growth factor
BLA	Biologics Licensing Application
BM	bone marrow
Bmax	maximum specific binding
BMP	bone morphogenic protein
BMS	Bristol-Myers Squibb
BMS	bare metal stent BS-seq bisulphite sequencing
BOC	boceprevir
CACNA1C	calcium voltage-dependent channel, L-type, α-1C subunit
CAD	coronary artery disease
cAMP	cyclic adenosine monophosphate
CARDS	Collaborative Atorvastatin Diabetes Study
CAV3	caveolin 3
CBZ	carbamazepine
CCR4	CC chemokine receptor 4
CD	cluster of differentiation
CDx	companion diagnostics
CE	capillary electrophoresis
CEIVD	European Communities *In Vitro* Diagnostic
CEVR	complete early virological response
CF	cultural fractionalization
CFTR	cystic fibrosis transmembrane conductance regulator
CFU-f	colony-forming unit of fibroblasts
cGMP	cyclic guanosine monophosphate
CHCA	α-cyano-4-hydroxycinnamic acid
CHCT	caffeine halothane contracture test
CHMP	Committee for Medicinal Products for Human Use
CID	collision induced dissociation
CK	creatine kinase
CK	cytokeratin
CNV	copy number variant
COMT	catechol O-methyltransferase
COX	ciclooxigenase; cyclo-oxygenase
C-path CAMD	Critical path coalition against major diseases
CpG/CG	phosphodiester bond between cytosine guanine

Handbook of Pharmacogenomics and Stratified Medicine. http://dx.doi.org/10.1016/B978-0-12-386882-4.00057-8
© 2014 Elsevier Inc. All rights reserved.

CPIC	Clinical Pharmacogenetics Implementation Consortium
CSF	cerebrospinal fluid
cTnI	cardiac troponins I
cTnT	cardiac troponins T
CYP	cytochrome P450; P-450
CYP450	cytochrome P450
CYP1A2	cytochrome P450, family 1, subfamily A, polypeptide 2
CYP2B6	cytochrome P450, family 2, subfamily B, polypeptide 6
CYP2C19	CYP2C19 cytochrome P450, family 2, subfamily C, polypeptide 19
CYP2C9	CYP2C9 cytochrome P450, family 2, subfamily C, polypeptide 9
CYP3A4	CYP3A4 cytochrome P450, family 3, subfamily A, polypeptide 4
DAA	direct-acting antiviral therapy
DAG	diacylglycerol
dbSNP	single-nucleotide polymorphism database
dbSNP	National Center for Biotechnology Information (NCBI) SNP database
DCS	diabetes care system
DDT	dichlorodiphenyltrichloroethane
DES	drug eluting stent
DHFR	dihydrofolate reductase
DHS	drug hypersensitivity syndrome
DHS	DNase I hypersensitivity sites
DIGE	difference gel electrophoresis
DIHS	drug-induced hypersensitivity syndrome
DILI	drug-induced liver injury
DISI	drug-induced skin injury
DMF	N,N-dimethyl formamide
DMEM	Dulbecco's modified essential medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DPD	dihydropyrimidine dehydrogenase
DPP	Diabetes Prevention Program
DRESS	drug reaction with eosinophilia and systemic symptoms
DVT	deep venous thrombosis
DWPG	The Royal Dutch Association for the Advancement of Pharmacy Pharmacogenetics Working Group DZ dizygotic
EASD	European Association for the Study of Diabetes
EBs	embryoid bodies
EBV	Epstein-Barr virus
EC	extracellular
ECC	excitation-contraction coupling
ECD	electron capture dissociation
ECG	electrocardiogram
ED_{50}	therapeutically effective dose in 50% of people given drug
EDTA	ethylenediaminetetraacetic acid
EED	embryonic ectoderm development
EET	epoxyeicosatrienoic acid
EF	ethnic fractionalization
EFV	efavirenz FDA Food and Drug Administration
EG	embryonic germ
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EM	extensive metabolizer
EMA	European Medicines Agency
EMF	electromagnetic field
EMHG	European Malignant Hyperthermia Group
emPAI	exponentially modified protein abundance index
ENCODE	Encyclopedia of DNA Elements
END2	endoderm-like cell 2
eNOS	endothelial nitric oxide synthase
ENU	*N*-ethyl-*N*-nitrosourea
eQTL	expression quantitative trait loci
ER	estrogen receptor
ES	embryonic stem
ESI	electrospray ionization
eSNP	expression single-nucleotide polymorphism
EST	expressed sequence tag
ETD	electron transfer dissociation
EZH2	enhancer of zeste
FASP	filter-aided sample preparation
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
FD	familial dysautonomia
FDA	U.S. Food and Drug Administration
FDR	false discovery rate
FGF	fibroblast growth factor
FGF20	fibroblast growth factor 20
FH	familial hypercholesterolemia
G6PD	glucose 6 dehydrogenase deficiency
GABA	gamma-aminobutyric acid
GEO	Gene Expression Omnibus
GLP	good laboratory practice
GO	Gene Ontology
Go-DARTs	Genetics of Diabetes Audit and Research Tayside
GPIa	glycoprotein Ia
GPIb/V/IX	glycoprotein Ib/V/IX complex
GP IIb/IIIa	glycoprotein IIb/IIIa
GPCR	G-protein coupled receptor
gPS	germline-derived pluripotent stem
GPVI	glycoprotein VI
GSCs	germline stem cells
GSST1	glutathione S-transferase theta 1
GST	glutathione S-transferase
GTEx	genotype tissue expression
GWAS	genome-wide association study
H3Africa	Human Heredity and Health for Africa project
HAART	highly active antiretroviral therapy
Hb	hemoglobin
HbA1c	glycated hemoglobin
HCD	high-energy collision-induced dissociation
HCV	hepatitis C virus
hERG	human *ether-à-go-go*–related gene
HER2	human epidermal growth factor receptor type2
HGF	hepatocyte growth factor
HIF-1α	hypoxia inducible factor 1α
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
HNF	hepatocyte nuclear factor
HPR	high platelet reactivity
HSCs	hematopoietic stem cell
HSS	hypersensitivity syndrome

HUCB-NSC	human umbilical cord blood–neural stem cell line
IC	imprinting center
IC	intracellular
ICAPS	International Consortium for Antihypertensives Pharmacogenomics Studies
ICAT	isotope-coded affinity tag
ICH	International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use
ICM	inner cell mass
ICPC	International Clopidogrel Pharmacogenomics Consortium
IEF	isoelectric focusing
IFNL	*interferon-lambda*
IGF1	insulin-like growth factor 1
IKBKAP	inhibitor of kappa light polypeptide gene enhancer in Bcells, kinase complex-associated protein
IKr	rapid delayed rectifier potassium current
IKs	slow delayed rectifier potassium current
ILSI/HESI	International Life Sciences Institute/Health and Environmental Sciences Institute
IM	intermediate metabolizer
IMDM	Iscove's modified Dulbecco's medium INDELs insertion and deletions
IND	investigational new drug IL interleukin
INH-DILI	isoniazid drug induced liver injury
INR	international normalized ratio
IPA	Ingenuity pathway analysis
iPS/C	induced pluripotent stem/cell
ISG	interferon-stimulated gene
ITP	inosine triphosphate
ITPA	inosine triphosphatase
IUPHAR	International Union of Basic and Clinical Pharmacology
IVCT	*in vitro* contracture test
IWPC	International Warfarin Pharmacogenetics Consortium
JLNS	Jervell and Lange-Nielsen syndrome
K_{ATP} channel	ATP-dependent potassium channel
KCNE1	potassium voltage-gated channel, ISK-related subfamily, member 1
KCNE2	potassium voltage-gated channel, ISK-related subfamily, member 2
KCNH2	potassium voltage-gated channel, subfamily H, member 2
KCNJ2	potassium channel, inwardly rectifying, subfamily J, member 2
KCNJ5	potassium channel, inwardly rectifying, subfamily J, member 5
KCNQ1	potassium voltage-gated channel, KQT-like subfamily, member 1
KD_{50}	agonist dissociation constant
KIM-1	kidney injury molecule-1
Klf4	Kruppel-like factor 4
KOSR	knockout serum replacement
KRAS	Kirsten rat sarcoma 2 viral oncogene homolog
LABA	long-acting $\beta 2$ adrenoceptor agonist
LC	liquid chromatography
LCL	lymphoblastoid cell line
LC-MS/MS	liquid chromatography–tandem mass spectrometry
LD	linkage disequilibrium
LDL	low-density lipoproteins
LEOPARD	lentigines, ECG anomalies, ocular hypertelorism, pulmonic stenosis, abnormal genitalia, retarded growth and sensorineural deafness
lncRNA	long noncoding RNA
LRRK2	leucine-rich repeat kinase 2
LTG	lamotrigine
MACE	major adverse cardiovascular events
MAF	minor allele frequency
MALDI	matrix-assisted laser desorption-ionization
MARS	Multiple Affinity Removal System
MATE1	multidrug and toxin extrusion 1
MCI	mild cognitive impairment
MDD	major depressive disorder
MDR1	multidrug resistance 1
MEA	microelectrode array
MeCP2	methyl CpG-binding protein 2
MEFC2C	myocyte enhancer factor 2
MEF/s	mouse embryonic fibroblast/s
MeHg	methylmercury
MethylC-seq	methylC sequencing
MELDI	material-enhanced laser desorption/ionization
MH	malignant hyperthermia
MHC	major histocompatibility complex
MHCI	major histocompatibility complex class I
MHCII	major histocompatibility complex class II
MHLW	Ministry of Health, Labor and Welfare
MiMI	Michigan Molecular Interactions
MI	myocardial infarction
miR-eQTL	microRNA quantitative trait loci
miRNA	microRNA
MMC	mitomycin C
MNNG	*N*-methyl-*N*-nitro-*N*-nitrosoguanidine
MPE	maculopapular eruption
MPE	maculopapular exanthema
MPP	1-methyl-4-phenylpyridinium
MRI	magnetic resonance image
MRM	multiple reaction monitoring
mRNA	messenger ribonucleic acid
MRE	microRNA recognition element
MSC	mesenchymal stem cell
MS	mass spectrometry
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MudPIT	multidimensional protein identification technology
MYH7	myosin-heavy polypeptide 7
MZ	monozygotic
N-2	neutropan-2
NADPH	nicotinamide adenine dinucleotide phosphate
NAMHG	North American Malignant Hyperthermia Group
NAT-2	(v) N-Acetyl Transferase-2
NAT	N-acetyltransferase
NCBI	National Center for Biotechnology Information

NCE	new chemical entity
NDA	new drug application
NEAA	nonessential amino acid
NGS	next-generation sequencing
NHS	National Health Service
NI	nucleotide inhibitor
NIH	National Institutes of Health
NKX2.5	NK2 Transcription Factor-Related
NLP	natural-language processing
NME	new molecular entity
NNI	nonnucleos (t) ide inhibitor
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNT	number needed to test/treat
NRTI	nucleoside reverse transcriptase inhibitor
NPV	negative predictive value
NSC	nearest shrunken centroid
NSTEMI	non ST elevation myocardial infarction
NT	neurotrophic
NVP	nevirapine
OATP1B1	organic anion–transporting polypeptide 1B1
OCP	oral contraceptive pill
OCT	organic cation transporter
Oct4/OCT4	octamer-binding protein 4
OR	odds ratio
OSM	oncostatin M
OXC	oxcarbazepine
P2Y$_{12}$	Purinergic receptor P2Y, G-protein coupled, 12
PAD	peripheral arterial disease
PAGE	polyacrylamide gel electrophoresis PGRN Pharmacogenomics Research Network
PAHs	polycyclic aromatic hydrocarbons
PANK	pantothenate kinase
PC	pheochromocytoma
PCA	principal component analysis
PcG	polycomb group
PCI	percutaneous coronary intervention
PCR	polymerase chain reaction
PCR-RFLP	polymerase chain reaction–restriction fragment-length polymorphism
PD	Parkinson's disease
PD	pharmacodynamics
pegIFN	pegylated interferon-α, RBV, ribavirin
pES	parthenogenetic embryonic stem
PET	positron emission tomography
PGC1β	peroxisome proliferator-activated receptor-γ co-activator 1β
PgWP	Pharmacogenomics Working Party
PGx	pharmacogenetics or pharmacogenomics
PharmGKB	Pharmacogenomics Knowledgebase
pI	isoelectric point
PI	protease inhibitor
PK	pharmacokinetics
PKL	Paxillin Kinase Linker
PM	poor metabolizer
PMDA	Pharmaceuticals and Medical Devices Agency
PMID	PubMed ID
POP	PMDA Omics Project
PPAR	peroxisome proliferator-activated receptor
PPARα	peroxisome proliferator-activated receptor α
PPARγ	Peroxisome proliferator-activated receptor γ
PPIs	Proton pump inhibitors
PPP	pentose phosphate pathway
PPV	positive predictive value
PRC	polycomb repressive complex
Pre-miRNA	precursor miRNA
Pri-miRNA	primary miRNA PTM post-translational modification
PRP LTA	light transmission aggregometry on platelet rich plasma
PSTC	Predictive Safety Testing Consortium
PTE	pulmonary thrombosis/thromboembolism
PXR	pregnane X receptor
QT	Q wave T wave
RA	retinoic acid
RC	Repressive Complex
RCT	randomized control trial
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
ROCK	rho-associated protein kinase rsID Reference SNP identification
RPA-1	Renal papillary antigen-1
rtPCR	real-time polymerase chain reaction
RVR	rapid virological response
RWS	Romano-Ward syndrome
RYR1	ryanodine receptor gene
SABA	short-acting β2 adrenoceptor agonist
SAND	Sp100, AIRE-1, NucP41/75, DEAF-1
SCN5A	sodium voltage-gated channel, type V, α-subunit
SCN4B	sodium voltage-gated channel, type IV, β subunit
SCZD	SNP single-nucleotide polymorphism
SELP	P-selectin
SILAC	stable isotope labeling by amino acids in cell culture
SJS	Stevens-Johnson syndrome
SMA	spinal muscular atrophy
SMN1	survival motor neuron 1
SMX	sulfamethoxazole SCAR severe cutaneous adverse reaction
SNP	single-nucleotide polymorphism
sSNP	splicing single-nucleotide polymorphism
SNTA1	syntrophin α1
Sox2/SOX2	SRY (sex-determining region Y)-box 2
SPA	sinapinic acid; SRM, selected reaction monitoring
sQTL	splicing quantitative trait loci
SSC	spermatogonial stem cell
SSEA4	stage-specific embryonic antigen 4
STEMI	ST elevation myocardial infarction
SULT	sulfotransferase SELDI surface-enhanced laser desorption-ionization
SUR1	sulfonylurea 1 receptor
SUZ12	Suppressor of Zeste 12
SVM,	support vector machine
SVR	sustained virological response
T2DM	type 2 diabetes mellitus
TB	tuberculosis

TCF7L2	transcription factor 7-like 2	UCSCs	umbilical cord stromal cells
TCR	T-cell receptor	UGT	uridine diphosphate glucuronosyltransferase
TEN	toxic epidermal necrolysis	UKPDS	UK Prospective Diabetes Study
TFF3	Trefoil factor 3	ULT	ULTRAPETALA
TGF-β	transforming growth factor-β	ULvWF	ultra large von Willebrand factor
TI	therapeutic Index	UM	ultrarapid metabolizer
TIA	transient ischemic attack	URM	ultrarapid metabolizer
TOF-MS	time-of-flight mass spectrometry	UTR	untranslated region
TOF	time-of-flight TSS transcription start site	VASP	platelet vasodilator-stimulated phosphoprotein
TPMT	thiopurine methyl S-transferase	VDR	vitamin D receptor
TRBP	TAR RNA binding protein	VIP	very important pharmacogene
TRIPOD	Troglitazone in the Prevention of Diabetes	VIP	variable importance in the projection
TrxG	trithorax group	VKOR	vitamin K epoxide reductase
TS	thymidylate synthase	VKORC	vitamin K epoxide reductase complex
T-tau	total tau	VKORC1	vitamin K epoxide reductase complex subunit 1
TVR	telaprevir	VKORC1	vitamin K 2,3-epoxide reductase C1 subunit
TxA2	thromboxane A2	VNTR	variable number of tandem repeats
UA	unstable angina	vWF	von Willebrand factor
UCB	umbilical cord blood	VXDS	voluntary exploratory data submission
UCMSC	umbilical cord matrix system cell	WHO	World Health Organization
UCPVC	umbilical cord perivascular cell	ZFN	zinc-finger nuclease

This page is intentionally left blank

Index

1D gel analysis, 152–153
1D-^{1}H-Nuclear Overhauser Effect Spectroscopy (NOESY) pulse, 189
2D-gel analysis, 153, 164–167
2D NOESY, 190
1200 Patients Project, University of Chicago, 926–927
α1 Acid glycoprotein (AAG), 890
α2-antiplasmin, 482
α-tubulin, 30
αβ immunotherapy, 571–573
β1-adrenergic receptor gene (ADRB1), 820, 1006–1007
β2 adrenoceptors, 374–375
β2-adrenergic receptor agonists, 636–640
β2-adrenergic receptor gene (ADRB2), 820
β3-adrenergic receptor (ADRB3), 1007–1008
β-amyloid buildup, 571
β-blockers, 408–410, 761–762
  *AD*RB1, 761
  *G*RK4, 761–762
β-glucosyltransferase (βGT), 61
β-tubulin, 30
μ receptor polymorphism, 829

A

Abacavir, 7–9, 418, 444–450
Abacavir hypersensitivity, 417–419
Abacavir hypersensitivity reactions (AHRs), 444–455
*AB*CB1, 357–358, 524, 658, 816, 1050
ABCB1 and ABCC2, 1003–1004
*ABCB11 (BS*EP), 473
ABCC1, 358, 1002–1003
*ABCC2 (MR*P2), 358, 473
*AB*CG2, 358–359, 661–662, 1004–1005
Abelcet, 881
Abetalipoproteinemia (ABL), 725
Absolute protein expression (APEX), 162
Absorption, 341–342
ACCE analytic framework:, 922
ACE, 763, 769
Acenocoumarol, 352
Acenocoumorol, 1054
Acetaminophen, 389, 409, 830
Acetone, 389
ACHIDO (Atorvastatin and Clopidogrel High Dose), 516
Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial, 741
Activated partial thromboplastin time (APTT), 480–481, 485
Activated protein C (APC), 377
Activity QTL (aQTL), 116
Acute coronary syndrome (ACS), 374, 487, 509
Acute generalized exanthematous pustulosis (AGEP), 417
Acute HCV infection, 868–869
Acute intermittent porphyria (AIP), 820
Acute lymphoblastic leukemia (ALL), 83, 345, 411, 945–946
Acute myeloid leukemia (AL), 82–83
Adalimumab (ADA), 668
ADAMTS-13, 518
Adaptive trial design, 317
*ADD*1, 769
Adenosine deaminase (ADA), 655
Adenosine diphosphate (ADP), 509–510
Adenosine monophosphate (AMP), 655
Admixture, 984
Ado-Trastuzumab Emtansine, 7–9
ADP-induced platelet aggregation, 525
ADP-inducible platelet reactivity, 517
*AD*RB1, 761
*AD*RB2, 636–638
  polymorphisms, 639
*ADRB2 Arg*16 and second-line therapy, 640–643
Adrenocorticotropic hormone (ACTH), 749–752
Adult stem cells, 208
Adverse drug reactions (ADRs), 3–4, 223–224, 378–379, 405, 936–937, 954
  burden of, 407
  classification, 405–407
  phase II metabolism, 410–414
    cisplatin, 411
    irinotecan, 412–413
    isoniazid, 411–412
    sulfamethoxazole (SMX), 412
    thiopurines, 410–411
    warfarin, 413–414
  phase I metabolism, 408–410
    CYP2C9, 410
    CYP2D6, 408
    tardive dyskinesia and antipsychotics, 410
  predisposition to, 407–408
    twin studies, 407–408
  type B adverse drug reactions, 414–424
    drug induced liver injury (DILI), 421–424
    drug-induced skin injury, 417–421
    malignant hyperthermia and anesthesia, 414–415
    statins and skeletal muscle toxicity, 415–417
Adverse Event Reporting System (AERS), 815
Afatinib, 18
Affected sib pair (ASP) study, 247
African populations, diversity of, 971–979
  and epigenetics, 972–973
  future perspectives, 992–994
    epigenetics, 992–993
    human heredity and health, 994
    from pharmacogenetics to pharmacogenomics, 993–994
  genetic variation and diversity, 979
  high genetic diversity, 979–985
    African populations’ admixture within and outside of Africa, 984
    dietary factors in population diversity and evolution, 984–985
    genetic structure, 980–982
    out-of-Africa population origin model, 982–984
  historical approaches, 973
  pharmacogenomics, 985–992
    pharmacogenetically relevant genes, variation in, 989–992
    pharmacogenetic diversity, 985–992
  population genetics, diversity, ancestry, and admixture, 979–985
  rich cultural and natural history, 973–978
    burden of disease, 977–978
    demographic, social, economic, and health-related indicators, diversity in, 976–978
    environmental health, 977
    life expectancy and age distribution, 977
    linguistic diversity, 973–974
    measuring population diversity, 974–976
    urban versus rural, 977
  and stratified medicine, 978–979
AGT, 763, 771
*AG*TR1, 763, 771
Airflow limitation, measurements of, 622
Airway hyper-responsiveness, 634–635
Airway inflammation, in asthma, 635–636
Airway remodeling, 623
Ala-Korpela group’s method, 185–186
Albendazole, 329–330
Alignment correction, 273
Allele frequency differentiation, in India, 1039–1045
Allergen exposure and asthma, 621
Allergic bronchopulmonary aspergillosis (ABPA), 621
Allergy and asthma, 631–632
Allopurinol, 12–13, 418, 452–453, 445–449

  hypersensitivity, 420
Allosteric modulation, 372
Alpha adducin (*A*DD1), 759
Alpha helices, 270
Alpha-1-antitrypsin (A1AT) deficiency, 224
Alpha1-antitrypsin–(AAT-) deficiency, 103
Alphahydroxymidazolam, 826–827
*Alternar*ia species, 621
Alzheimer's disease (AD), 563
  biomarkers and comorbidity, 569–571
    body fluids, biochemistry of, 570
    genomics and proteomics, 570–571
    neuropathology, 569
    neurophysiology, 569–570
    structural and functional neuroimaging, 569
  future perspective, 584–585
  genes associated with, 586–608
  genomics of, 565
  immunotherapy, 571–573
  pathogenic events, 565–569
    cerebrovascular dysfunction, 567
    cholesterol and lipid metabolism dysfunction, 568
    epigenetic phenomena, 567
    genomic defects, 566–567
    neuroinflammation and immunopathology, 568
    neuronal apoptosis, 567–568
    neurotoxic factors, 568
    neurotransmitter deficits, 568
    oxidative stress, 568
    phenotypic expression of amyloid deposits and neurofibrillary tangles, 567
  pharmacogenomics, 573–584
    antidementia drugs, pharmacogenomics of, 581–582
    CNS drugs, genes involved in the mechanism of action of, 574
    dementia, pleiotropic activity of *A*POE in, 580–581
    drug metabolism, genes involved in, 574–580
    drug transporters, 580
    multifactorial therapy, 582–584
    pharmacogenomics, 573–574
  therapeutic strategies, 571–573
Ambisome, 881
American Academy of Pediatrics (AAP), 936
American Society of Clinical Oncology (ASCO), 946
American Standard Code for Information Interchange (ASCII), 281
Aminoglycosides, 853
Aminoimidazole carboxamide ribonucleotide (AICAR), 655
Aminopenicillins, 418, 445–449
Aminotransferase (ALT), 828
Amitriptyline, 353–354
Amniotic fluid (AF) stem cells, 209
Amniotic membrane stem cells, 209
Amoxicillin clavulanate (AC), 12–13, 422–423, 441–442, 445–449, 453–455
Amphotericin B, 890–891, 893, 895
AmpliChip$^{®}$ CYP450 Test, 534
AmpliChip™, 944–945
Analgesics, 1056
Analphalipoproteinemia, 726
Analysis of variance (ANOVA), 135, 250–252
Anesthesia, pharmacogenomics in, 815
  metabolism, 816–820
    modified drug metabolism, 818
    non-CYP enzymes, 818–819
    phase I enzymes, 816–818
    phase II enzymes, 819–820
  pharmacodynamics, 820–830
    butyrylcholinesterase, abnormalities in, 821–822
    defective enzymes, 820–821
    inhalational anesthetic agents, 827–828
    malignant hyperthermia (MH), 822–825
    neuromuscular blocking drugs (NMBDs), 828
    nonsteroidal anti-inflammatory drugs (NSAIDs), 830
    opioids, 828–830
    specific drugs, 825–827
  transport proteins, 815–816
Angelman syndrome (AS), 65
Angiotensin II, 763
Angiotensin receptor 1 (*AG*TR1) gene, 395–396
Angiotensin receptor blockers (ARBs), 763–764
  ACE, 763
  *A*GT, 763
  *AG*TR1, 763
  GWAS findings, 763–764
Angiotensin-converting enzyme (ACE) inhibitors, 12–13, 369, 763–764, 818, 1051–1052
  ACE, 763
  *A*GT, 763
  *AG*TR1, 763
  GWAS findings, 763–764
Anglo-Scandinavian Cardiac Outcomes Trial–Blood Pressure Lowering Arm (ASCOT-BPLA), 754
Anidulafungin, 893
Animal experimentation, ethical issues of, 80
  reduction, 80
  refinement, 80
  replacement, 80
Animal models in pharmacogenomics, 73
  applications of in vivo models, 81–82
    identification of eQTLs, 81–82
    quantitative trait mapping, 81
  caution and limitations, 80–81
  ethical issues of animal experimentation, 80
    reduction, 80
    refinement, 80
    replacement, 80
  genetic/genomic resources, 79–80
  nonrodent models of human disease, 79
  rodent models, 74–79
    collaborative cross (CC), 75–76
    congenic strains, 76–77
    consomic strains, 77
    heterogeneous stocks, 75
    inbred strains, 75
    mouse models, 74
    outbred rodent stocks, 74–75
    rat models, 74
    recombinant-inbred strains (RIS), 76
    transgenic rodent models, 77–79
  rodent models, 82–85
    anxiety disorders, 82
    bipolar disorders, 82
    pharmacodynamic/pharmacogenomic modeling of insulin resistance, 84–85
    pharmacogenomics and PAH, 84
    pharmacogenomics of pain and analgesia, 83–84
    pharmacogenomics using xenograft mouse models of acute leukemia, 82–83
Animal–plant warfare, 995
Anthracycline, 1053
Antiarrhythmic agents, 791
Anticancer drugs, 1052–1053
  anthracycline, 1053
  letrozole, 1052–1053
  tamoxifen, 1052
  thiopurines, 1053
Anticoagulants, oral, 488–494
  direct thrombin inhibitors, 491–493
    dabigatran, monitoring the anticoagulant effect of, 492–493
    interactions, 491–492
    nonhaemorrhagic side effects, 493
    pharmacokinetics, 491
    pharmacology, 491
  factor Xa inhibitors, 493–494
    bleeding and reversal of anticoagulant effect, 494
    interactions, 493
    mechanism of action, 493
    monitoring of anticoagulant effect, 493–494
    nonhemorrhagic adverse effects, 494
    pharmacokinetics, 493
  vitamin K antagonists (VKAs), 489–491
    interactions, 489–490
    management of bleeding and reversal, 490–491
    mechanism of action, 489
    monitoring the anticoagulant effect of warfarin, 490
    nonhaemorrhagic adverse effects, 491
    pharmacokinetics, 489
Anticoagulants, parenteral, 484–488
  danaparoid, 487
  direct thrombin inhibitors, 487–488
    argatroban, 488
    bivalirudin, 488
  fondaparinux, 487
    adverse effects, 487
    dosing and reversal, 487
    pharmacokinetics, 487
    structure and mechanism of action, 487
  low-molecular-weight heparin (LMWH), 486–487
  unfractionated heparin (UFH), 484–486
    adverse effects, 485–486
    monitoring, 485
    pharmacogenomics, 486
    pharmacokinetics, 485
    reversal, 486

    structure and mechanism of action, 484–485
Anticoagulation, indications for, 482–484
  atrial fibrillation (AF), stroke prevention in, 483–484
  venous thromboembolism, 482–483
    epidemiology, 483
    management, 483
    pathophysiology, 483
Anticonvulsants, 1055
Antidementia drugs, pharmacogenomic profile of, 575–576
Antidepressant drugs (AD), 543–544, 991–992
  candidate gene studies, 544–553
    pharmacodynamics, 544–552
    pharmacokinetics, 552–553
  genome-wide approaches, 553–555
    genome-wide association studies (GWAS), 553–554
    pathway analysis, 554–555
Antidiabetic drugs
  genetic basis of diabetes, 683–684
  genetic variants associated with, 685–709
    biguanides (metformin), 685–703
    nonsulfonylurea insulin secretagogues (glinides), 706–707
    sulfonylureas, 703–706
    thiazolidinediones, 707–709
  management of, 684–685
  pharmacogenomics, 685
  stratified medicines, 685
  stratified oral antidiabetic medicines, 709–710
Antiemetics, 830
Antiepileptic drugs (AEDs), 421
Antifolates, 1056
Antifungal agents, 879–886
  cellular targets of, 880
  echinocandins, 884–885
  flucytosine, 885
  future perspectives, 896
  genetic variability associated with, 887–895
    absorption, 887–890
    distribution, 890–892
    elimination, 895
    metabolism, 892–895
  in vitro spectrum of, 880
  polyenes, 879–881
  terbinafine, 885–886
  triazoles, 881–884
    fluconazole, 882
    itraconazole, 882–883
    posaconazole, 883–884
    voriconazole, 883
Antigen-presenting cells (APCs), 103, 468
Antihistamines, 793
Antihypertensives, 754–758, 764–771
  clinical trials, 764
  genes with strong evidence of association with, 764–769
    ACE, 769
    *A*DD1, 769
    *AD*RB1, 764–765
    AGT, 771
    *AG*TR1, 771
    *CACN*A1C, 766
    *CAC*NB2, 765–766
    fibrinogen-beta (FGB) gene, 766–768
    gene panels, 771
    genetic risk scores, 768–769
    *KCN*MB1, 769–770
    liver X receptor-α (LXRA), 770
    *M*MP9 and *MM*P12, 770
    *NEDD*4L, 765
    *N*OS3, 770
    *N*PPA, 770
    *R*YR3, 770–771
Anti-IgE, 633–634
Anti-IL5, 634
Antiplatelet drugs, 373–376, 511
Antipsychotic medications, 793
Antipsychotics, 410, 991–992
  PGX of, 1029–1030
Antiretrovirals, pharmacokinetics of, 1024–1025
Antisense oligonucleotide (ASO), 387, 737–738
Anti-TB drugs, 850
Antituberculosis, 422
  medications, 420–421
Anxiety disorders, 82
Apixaban, 488, 492
Apo B/Apo A-1, 722
APOA-1 mutations, 725–726
*A*POE-related AD phenotypes, 581
*A*POE-related blood lipid response to sardilipin, 583–584
*A*POE-related changes in blood pressure values, 582–583
*A*POE-related cognitive function changes, 582
Apolipoprotein E (Apo E), 733
Apolipoprotein L1 (APOL1), 903–905
Apolipoproteins, 717, 721–722
Application-independent data analysis, 129–131
  assembly through read mapping, 130–131
  base calling, 129
  read assembly, 130
    *De N*ovo Assembly, 130
  read quality control, 129–130
  sequence and read count, 131
Aquaglyceroporin-type transporter (*A*QP2), 910–911
*AR*G1 polymorphisms, 639
Argatroban, 488
Aripiprazole, 7–9
ARITMO (Arrhythmogenic Potential of Drugs) project, 802
Aromatase inhibitors, 12–13
Arsenic Trioxide, 7–9
Aryl hydrocarbon receptor (AhR), 391–392
Aryl hydrocarbon receptor nuclear translocator (ARNT), 392
Arylsulfatase, 893
Aspergillus, 883
Aspergillus terreus, 729
Aspirin, 12–13, 445–449
"Aspirin resistance", 373–374
Assembly through read mapping, 130–131
Assent in pediatric research, 940–943
Association studies, 247–256
  association study designs, 247–249
    candidate-gene studies, 247–249
    genome-wide association studies (GWAS), 249
  correction for multiple testing, 253–254
    population stratification and admixture, 253–254
  designs, 247–249
    candidate-gene studies, 247–249
    genome-wide association studies (GWAS), 249
  fine mapping and resequencing, 255–256
  GWA studies, reporting of, 256
  imputation, 254–255
  meta-analysis, 255
  population-based study designs, 249–253
    study power and coverage, 252–253
    tests of association, 250–252
  validation and replication, 255
Asthma, 617–618
  clinical aspects, 618–623
    associated comorbidities and confounders, 619
    diagnosis, 621–623
    exacerbations, 620–621
    symptoms, 618–619
    treatment, 619–620
  cost, 618
  definition, 618
  drug development in, 645–647
    future directions, 647
    screening, stratification, personalization, 646–647
  genome-wide association studies (GWAS) in, 628–629
  incidence/prevalence, 618
  management of, 629–631
    bronchodilators, 629
    inhaled corticosteroids (ICS), 629–630
    pharmacological treatments and alternative therapies, 630–631
  pathophysiology, 623–629
    biological basis, 623
    genetic basis, 623–629
  specific treatment for, 631
  stratified medicine for, 631–645
    *ADRB2 Arg*16 and second-line therapy, 640–643
    clinicopathophysiological measures, 634–636
    corticosteroid pharmacogenetics, 643–645
    heterogeneity in, 631–632
    nongenetic approaches to patient stratification, 633–634
    pharmacogenetic approaches, 636–640
    problems with current therapy, 632–633
  targets of current asthma medication., 641
Ataxia telangiectasia mutated (ATM), 702–703
Atenolol, 756
*ATIC/CC*ND1, 660–661
Atmospheric-pressure chemical ionization (APCI), 196
Atomoxetine, 7–9
Atorvastatin, 7–9
ATP-binding cassette (ABC) transporters, 342, 655, 726, 733, 815–816, 1003–1004

Atrial fibrillation (AF), 791
  stroke prevention in, 483–484
Autonomic nervous system, hypertension and, 753
"Autonomy", 935
Autosomal dominant hypercholesterolemia (ADH), 725
Autosomal recessive hypercholesterolemia (ARH), 725
Autosomes, 33
Avon Longitudinal Study of Parents and Children (ALSPC), 95
Azathioprine (AZA), 7–9, 663–665
  glutathione S transferases (GSTs), 664
  *I*TPA, 664
  pharmacogenetics of, 665
  purine pathway enzymes, 664–665
  *T*PMT, 664

B

B cells, EBV transformation effects on, 95–96
  gene expression changes, 96
  immunophenotypes, 95–96
Backbone generation, 273
Backcrossing, 76
Bacterial agents, of asthma, 621
Baicalin, 1011
Barbiturates, 821
Bare-metal stents (BMS), 511
Base calling, 129
Bassen-Kornzweig syndrome. *See* Abetalipoproteinemia (ABL)
*B*CHE allele, 10
Beatson, George, 5–10
Beckwith-Wiedemann syndrome, 65
Benzodiazepeines (BDZs), 825–827
Benzodiazepine, 333, 937
Beryllium, 445–449
Best Pharmaceuticals for Children Act (BPCA), 938
Beta sheets, 270–271
Beta-1-selective metoprolol, 351
Beta-blockers, 786–787
Bevacizumab, 12–13
Biguanides (metformin), 685–703
  ataxia telangiectasia mutated: gene polymorphisms and metformin response, 702–703
  MATE1:SLC47A1 gene polymorphisms and metformin response, 702
  OCT2:SLC22A2 gene polymorphisms and metformin response, 702
  *SLC22*A1 gene polymorphisms and metformin response, 687–702
Bile acid sequestrants (BASs), 736–737
Biliary excretion, 348
Binary fission, 30–31
Binning, 198–199
Biobanks, LCL generation in, 92–93
  blood cell banking for LCL generation, 92
  current biobanks, 92
  quality control of LCLs, 92–93
Biogenesis, 385–386
Bioinformatics, 259
  bioinformatics databases, 276–284
    annotation, 280
    common file formats, 280–284
    gene expression, 277–278
    general sequence databases, 276–277
    noncoding sequences and epigenomics, 278–279
    pharmacogenomics, 280
    proteins, metabolites, and pathways, 279–280
  de novo structure prediction, 276
  fold recognition, 275
  homology modeling, 272–275
    alignment correction, 273
    backbone generation, 273
    loop modeling, 273–274
    model optimization, 275
    model validation, 275
    side-chain modeling, 274–275
    template recognition and initial alignment, 273
  molecular evolution and phylogenetics, 264–269
    methods of tree inference, 267–269
    principles of tree inference, 266–267
    tree terminology, 265–266
  nucleic acids and proteins, analysis of sequences of, 260–264
    comparison and alignment of two sequences, 261–263
    database searching, 264
    dot matrix comparison of two sequences, 263
    multiple sequence alignment, 263–264
    small motifs, 261
  protein data bank file format, 271–272
  protein modeling, 272
  protein stability, 269–271
    allowed main-chain conformation, 270
    alpha helices, 270
    beta sheets, 270–271
    permitted torsion angles, 270
    turns and loops, 271
  threading, 275–276
Biologic DMARDS, pharmacogenetics of, 668–674
  rituximab, 673–674
  tocilizumab, 674
  tumor necrosis factor antagonists, 668–673
    cytokines, 672
    Fcγ receptor variants, 672
    Genome-wide association studies (GWAS), 672–673
    MHC gene polymorphisms, 671
    MHC Microsatellites, 671
    RA risk variants, 671–672
    signaling pathways, 672
    TNF gene polymorphisms, 668–670
    TNF receptor polymorphisms, 670–671
    toll-like receptors, 672
Biological resource centers (BRCs), 92
Biomarker, 13–16
Biomarker-guided dosing, 313
BioMe Biobank Program, 927
Biomedical research with human subjects, ethics of, 934
Biorepository and privacy issues, 813
Bipolar disorders, 82
Bisulphite treatment, 61
Bivalirudin, 488
BLAST, 264
Blood, 150
Blood pressure (BP), 747–748
  heritability of, 748–749
Blood pressure response to antihypertensive drugs, 758–764
  β-blockers, 761–762
  ACE inhibitors and ARBs, 763–764
  calcium channel blockers (CCB), 762–763
  thiazide diuretics, 758–761
Blood–brain barrier (BBB), 342, 567
BLOSSUM matrices, 260
B-lymphocyte–stimulating factor (BLyS), 673
*B*MP7, 553–554
Boceprevir, 7–9
*B*RAF gene, 376
Brain-derived neurotrophic factor (*B*DNF), 551
Brazil, pharmacogenomics in, 1015
  diversity among Brazilians, 1017–1021
  genetic clusters, 1018
  PGx clinical trials, 1021
    of antipsychotics in schizophrenic patients, 1030–1032
    antiretrovirals, 1024–1025
    of immunosuppressants in renal transplant patients, 1028–1030
    L-thyroxine in differentiated thyroid cancer, 1028
    NSAIDs, 1021–1024
    warfarin-dosing algorithms, 1026–1028
  Refargen, 1015–1029
Breast cancer resistance protein (BCRP), 888
Breeding strategies, 77
Brentuximab Vedotin, 7–9
Bromobenzene, 389
Bronchodilators, 629
BS-seq, 136
Bucketing. *See* Binning
Bupropion, 353
Burrow Wheeler transform (BWT), 130–131
Busulfan, 7–9
Butrylcholinesterase (BChE), 819
Butyrylcholinesterase (BChe), abnormalities in, 821–822
  qualitative variants, 821–822
  quantitative variants, 821

C

^{13}C signals, 187
Ca^{2+} transport, 822
5caC (5-carboxylcytosine), 61
*CACNA*1C, 766–767
*CAC*NB2, 765–767
Caffeine-halothane contracture test (CHCT), 414–415
Calcineurin inhibitors, 1028
Calcium channel blockers (CCB), 762–763
  *KCN*MB1, 762–763

Calumenin, 502–503
Canadian Pharmacogenomics Network for Drug Safety (CPNDS), 425
Cancer markers, 167
Cancer therapy, use of stem cells for, 226
Cancer-related P450s, regulation of, 391–394
  CYP19A1, 393–394
  CYP1A1, 391–392
  CYP1B1, 392
  CYP2A, 392–393
  CYP2J2, 393
  CYP24A1, 394
Candesartan, 12–13
Candida, 883
Candidate-gene studies, 247–249
Capecitabine, 7–9
Capillary electrophoresis (CE), 155
Capillary tube centrifugation technique (CTC), 908
Captopril, 756
Carbamazepine, 7–10, 12–13, 353–354, 418, 421, 450–452, 445–449, 1008, 1010
  hypersensitivity, 419–420
Carbon tetrachloride, 389
Carboplatin, 12–13
Carboxylesterases (CESs), 818–819
Card agglutination test for trypanosomiasis (CATT), 907
Cardiac dysrhythmias, 824
Cardiac myocyte ion currents and action potentials, 781–784
Cardiotoxicity studies, use of stem cells in, 221–223
Cardiovascular disease (CVD), 167–170, 715–716, 718–722
Cardiovascular drugs, 1053–1054
  acenocoumorol, 1054
  clopidogrel, 1053
  statins, 1054
  warfarin, 1053–1054
Cardiovascular risk, 722
Carisoprodol, 7–9
Carotid intima-media thickness (cIMT), 735–736, 740
Carr-Purcell-Meiboom-Gill (CPMG) pulse program, 189
Cartridge, 531
Carvedilol, 7–9
"Cascade screening", 786
Case-cohort study, 319
Case-control design, 101–102
Caspofungin, 891, 893, 895
Catechol O-methyltransferase (COMT), 112, 114, 818, 893, 1049
  *C*OMT gene, 547
Caution and limitations, 80–81
CCR4-NOT complex, 49–50
CD19-immunoglogulin co-ligation method, 95
Cedars-Sinai Medical Center, 95
Celecoxib, 7–9
Cell, overview of, 29–30
  membrane-bound organelles in eukaryotic cells, 30
  nonmembranous organelles in eukaryotic cells, 30
  nucleus and chromosomes, 29–30
Cell cycle, 30–33
Cell reproduction, 30–33
Cell-based assays, 218
Cell-based hemostasis model, 482
CE-MS, 153, 155–156
Centers for Disease Control and Prevention (CDC) Office, 924
Centromere, 31–32
CEPH collections, 93–94
Cerebral amyloid angiopathy (CAA), 567
Cerebrospinal fluid (CSF), 570
Cerebrovascular dysfunction
  in Alzheimer's disease, 567
Certolizumab pegol (CZP), 668
*C*ES1 gene, 492
Cetuximab, 7–9, 18
Cevimeline, 7–9
CFTR protein, 376
Chain termination method, 125–126
Chemical Ionization (CI), 197
Chemical shift, intensity, and coupling patterns, 198
Children
  tuberculosis among, 838, 844–845
    development and natural history, 844
    diagnosis and treatment, 844–845
Children's Oncology Group (COG), 941
China, pharmacogenomics in, 999
  CYP2C19-mediated herb–drug interaction, 1011
  drug-metabolizing enzymes, 999–1002
    CYP2C19, 999–1001
    CYP2D6, 1001
    CYP3A, 1001
    histamine N-methyltransferase (HNMT), 1001–1002
  drug receptors, 1006–1009
    GABA Receptor, 1008–1009
    peroxisome proliferator-activated receptor γ2 (PPAR-γ2), 1008
    β1-adrenergic receptor, 1006–1007
    β3-adrenergic receptor (ADRB3), 1007–1008
  drug transporters, 1002–1006
    ABCB1 and ABCC2, 1003–1004
    ABCC1, 1002–1003
    ABCG2, 1004–1005
    SLC30A8, 1006
    SLCO1B1, 1005–1006
  future perspectives, 1012
  ion channels, 1009–1011
    KCNH2, 1009–1010
    KCNJ11, 1010
    KCNQ1, 1010
    SCN1A, 1010–1011
  OATP1B1-mediated herb–drug interaction, 1011–1012
  translational approaches, 1012
ChIP sequencing (ChIP-seq) experiments, 129
ChIP-BS, 65
ChIP-seq, 135–136
*C*HL1 protein, 101–102
Chlordiazepoxide and Amitriptyline, 7–9
Chloroform, 389
Chloroquine, 7–9
Cholesterol
  metabolism dysfunction, 568
  synthesis of, 716–717
Cholesterol and Recurrent Events (CARE) trial, 720
Cholesterol ester transfer protein (CETP), 718, 733, 738–739
  deficiency, 726
Cholesterol synthesis, 716–717
Cholesterol Treatment Trialists' (CTT) Collaboration, 720–721
Chromatin, 29–30, 62
Chromatin immunoprecipitation (ChIP), 64, 213, 215
Chromatin modifications and gene expression, 50–51
Chromatin structure, 42–43
Chromosome imprinting, 59
Cimetidine, 817
Cisapride, 794
Cis-eQTL, 115
Cisplatin, 7–9, 12–13, 360, 411
Citalopram, 7–9, 12–13
Clarithromycin, 794
Class III agents, 791–793
Clearance, 349
CLIA-certified laboratories, 924
Clinical and research ethics
  historical perspective, 933–935
Clinical decision support (CDS) systems, 925–926
Clinical genetic testing, 921
Clinical pharmacogenetics and implementation, 922–924
  education, guidelines, and recommendations, 922–924
  pharmacogenetic testing
    regulation, 924
    turnaround time, 924
  validity and utility, 922
Clinical Pharmacogenetics Implementation Consortium (CPIC), 289, 296–297, 301, 922–923, 944
Clinical Risk Assessment Engine (CRAE), 927
Clinical utility of a test, 18
Clinical validity, 18, 922
CLIPMERGE–PGx program, 927
CLIPMERGE platform, 927
Clobazam, 7–9
Clomipramine, 7–9
Clonal amplification, 127
Clopidogrel, 7–9, 12–13, 509–514, 1053
  Alternative Thienopyridine-Class Chemical Compounds, 513–514
    prasugrel, 514
    ticagrelor, 514
  current status, 527–535
    pharmacogenetic testing, 529–535
    screening, stratification, and personalization, 527–529
  future directions, 535–536
  genetic dissection of clopidogrel respons, 518–527

    - _CYP2C19_*2 and adverse cardiovascular events, 523–526
    - _CYP2C19_*2 and high on-clopidogrel platelet reactivity, 519–523
    - genetic determinants, 518–519
    - investigated genetic variants, 526–527
  - metabolism, 511–513
  - PGxpress page for, 304
  - platelets, atherothrombotic diseases, and antiplatelet therapy, 509–511
  - response variation, 514–518
    - high on-clopidogrel platelet reactivity, 514–518
- Clozapine, 7–9, 445–449
- cLQTS-associated gene, 785
- ClustalX, 263
- Clustering methods, 269
- CNS disorders, 564–565
- CNS drugs, genes involved in the mechanism of action of, 574
- Cochran-Armitage trend test, 252
- Codeine, 7–9, 409–410, 829, 937
- Co-developing a drug, 19–21
- Collaborative cross (CC), 75–76
- Collagen (*COL22*A1) gene, 639
- College of American Pathologists (CAP), 924
- Collisionally activated dissociation (CAD), 195
- Collisionally induced dissociation (CID), 195
- Colony-forming unit of fibroblasts (CFU-f) assay, 219
- Companion diagnostic (CDx), 16–17, 954, 963
- Companion Diagnostics Project Team, 955
- Competence, 938
- Competitive antagonism, 371–372
- Complete early virological response (cEVR), 867
- Congenic strains, 76–77
- Congenital adrenal hyperplasia (CAH), 749–752
- Congenital long-QT syndrome (cLQTS), 784–787
  - clinical assessment and management, 786–787
  - in drug-induced LQTS, 795–796
  - rare mutations in, 800
- “Consent” in pediatrics, 938–940
- Consomic strains, 77
- Constitutive androstane receptor (CAR), 852
- Constitutive heterochromatin, 43
- Convergent functional genomics (CFG) approach, 82
- “Convergent functional genomics”, 82
- Copy number variants (CNVs), 54, 98, 119–120, 239–241
  - mechanisms of CNV formation, 54
  - population genetics of, 54
  - relationship of, and transcriptional regulation, 54
- Coronary artery disease (CAD)
  - clopidogrel therapy, 511, 515–516
- Coronary Drug Project (CDP), 740
- Corrected QT interval (QTc), 781
  - drug-induced variation in, 787–796
    - congenital LQTS genes in, 795–796
    - detecting drug-induced effects on repolarization, 795
    - human ether-à-go-go gene channel, 788
    - *KC*NH2 polymorphisms and disease susceptibility, 796
    - mechanisms of proarrhythmia in QT prolongation, 788–790
    - pharmacokinetic and pharmacodynamic factors, 794–795
    - prevention, monitoring and treatment of LQTs, 795
    - QTc prolongation–inducing drugs, 791–794
    - reduced repolarization reserve, 790–791
    - repolarization heterogeneity, 788
  - in new drug discovery and drug safety, 797–798
- Correction for multiple testing, 253–254
  - population stratification and admixture, 253–254
- Correlaction Spectrsocopy (COSY), 189–190
- Corticosteroid pharmacogenetics, 643–645
  - combination therapy, 644–645
  - corticotrophin-releasing hormone receptor 1 (*CR*HR1) gene, 618–619
  - dual-specificity phosphatase 1 (*DU*SP1) gene, 644
  - low-affinity IgE receptor gene, 644
  - stress-induced phosphoprotein 1 gene (*ST*IP1), 644
  - T box 21 and neurokinin 2 receptor genes, 619
- Corticosteroids (CS), 84, 654–655
  - and anti-IL13, 634
- Corticotrophin-releasing hormone receptor 1 (*CR*HR1) gene, 618–619
- Co-trimoxazole, 418
- Coumarin anticoagulant, 497–498, 501
- Coverage, 847
- CPD-0802, 913
- CpG islands, 58
- Creatine kinase (CK), 415–416
- *CR*EB gene (cAMP-responsive element binding protein), 552
- Cre-LoxP Technology, 77
- Crizotinib, 7–9, 18, 376
- *Cryptococcus*, 883
- Cultural fractionalization (CF) index, 974–976
- CXCL10, 909
- CXCL8, 909
- Cyclic guanosine monophosphate (cGMP), 368
- Cyclophosphamide, 353
- Cyclosporin A
- Cyclosporine, 354, 1001
- Cyclosporine A (CSA), 91, 1028, 1055–1056
- CYP, 667
- CYP 1 gene, 329–330
- CYP 2 gene, 335
- CYP clustering, 579–580
- CYP1A1, 391–392, 1039
- CYP1A2, 388, 1039
  - phenotypes, 330
- CYP1B1, 392
- CYP24A1, 394
- CYP2A, 392–393
- CYP2A6, 335, 388, 1039
- CYP2B6, 335, 388, 989–991
- CYP2B6*6 carriers, 800
- *CY*P2C genes, 389
- *CYP2*C19, 10, 333, 420–421, 513, 524, 578, 999–1001, 1039
- CYP2C19 genetic variants, 520
- *CYP2C*19 genotyping strategy, 925
- _CYP2C19_*2, 519, 924
  - and adverse cardiovascular events, 523–526
  - genetic testing, 529
  - genotyping technologies, 534
  - and high on-clopidogrel platelet reactivity, 519–523
- CYP2C19-mediated herb–drug interaction, 1011
- CYP2C8, 389, 1039
- CYP2C8*3 allele, 331
- CYP2C8/9/19, 330–333
- CYP2C9, 388, 410, 420–421, 501–502
  - phenytoin, 410
- CYP2C9 alleles, 331–333, 578, 992
- *CYP2*C9 genotype, 10
- CYP2C9 polymorphisms, 1022
- CYP2D6, 5–10, 333, 408, 991–992, 1001, 1039
  - codeine, 409–410
  - β-blockers, 408–410
- *CYP2*D6, 112, 114, 574–578
- CYP2E1, 112–114, 336, 388–390, 473, 1039
- *CYP2*E1 polymorphisms, 851
- *CYP2*E1 variants, 112–113
- CYP2J2, 393
- CYP3A, 1001
- CYP3A4, 112–114, 333–335, 388, 390–391, 1039
  - isoenzyme, 731
- *CYP3A*4/5, 578–579
- CYP3A5, 112–114, 333–335, 388, 989–990, 1047
- *CYP3*A5, 297–299, 524, 731–732
- CYP3A7, 388
- CYP19A1, 393–394, 1051
- CYP450 2D6, 324
- CYP450 and ethnicity, 323
  - catalytic cycle, 325
  - future perspectives, 337
  - importance of race/ethnicity, 326–327
  - induction, 325–326
  - phenotyping, 327–329
  - variation, in different racial/ethnic populations, 329–330
    - CYP 1, 329–330
    - CYP2A6, 335
    - CYP2B6, 335
    - CYP2C8/9/19, 330–333
    - CYP2D6, 333
    - CYP2E1, 336
    - CYP 2 gene, 335
    - CYP3A4 and CYP3A5, 333–335
- CYP450 enzymatic pathways, 729
- CYP450 inhibitors, 346
- CYP450 isoenzymes, 516–517
- CYP4F2, 502
- Cysteinyl leukotriene receptor 1 and 2 genes (*CYSLTR1/CYSL*TR2), 642

Cystic fibrosis, 376
Cytochrome p450 (CYP) genes, 408, 471–472
Cytochrome P450 (CYP450) enzymes, 350–354, 385, 512–513, 552–553, 829, 850, 881–882, 892, 937, 985–986, 999
  CYP2A6, 353
  CYP2B6, 353
  CYP2C19, 352–353
  CYP2C8, 353–354
  CYP2C9, 351–352
  CYP2D6, 350–351
  CYP3A4/5, 354
Cytochrome P450 2C8 (CYP2C8), 706
Cytochrome P450 2D6 (CYP2D6), 1046
Cytochrome P450 2E1 (CYP2E1), 850
Cytochrome P450 system (CYP), 662–663, 816
  CYP2C, 817–818
  CYP2C19, 818
  CYP2D, 817
  CYP2D9, 817
  CYP2E1, 817
  CYP3A, 817
  CYP3A5, 817
Cytokine analysis, 905
Cytokinesis, 34
Cytoplasm, 29
Cytotoxic T lymphocytes (CTLs), 103

D

Dabigatran, 12–13, 488, 492
  bleeding and reversal of the anticoagulant effect of, 492–493
  etexilate. *See* Direct thrombin inhibitors
  monitoring the anticoagulant effect of, 492–493
Dabrafenib, 18
Dalcetrapib, 739
Danaparoid, 487
Dapsone, 7–9, 445–449
Dasatinib, 7–9
Data analysis, 163–167
  biomarker pattern analysis, 164–167
  data integration and OMICS analysis, 167
  downstream processing, 163–164
    identification, 164
    normalization, 163
    statistical analysis, 163–164
Daunorubicin HCl, 357–358
D-dopachrome tautomerase (DDT), 116
De Bruijn graphs, 130
*De N*ovo Assembly, 130
*De n*ovo genome sequencing, 131
De novo structure prediction, 276
Declaration of Helsinki, 933
Deep vein thrombosis (DVT), 483
Defective enzymes and anesthesia, 820–821
Dementia
  personalized medicine for, 564–565
  pleiotropic activity of *A*POE in, 580–581
Denileukin Diftitox, 7–9
Deoxycytidine kinase (dCK), 808–809
Deoxythymidine monophosphate (dTMP), 659
Deoxyuridine monophosphate (dUMP), 659
Depolarizations, 779–780
Depolarizing muscle relaxants (DMRs), 828
Dermatophytosis, 886
Desensitization and tachyphylaxis, 372–373
Desipramine, 7–9
Deuterium oxide, 189
Developing deep vein thrombosis (DVT), 377
Dexlansoprazole, 7–9
Dextromethorphan and Quinidine, 7–9
*DHFR*P1 pseudogene, 396
*DH*ODH, 661
Diabetes, 683, 730–731
  genetic basis of, 683–684
  management of, 684–685
  metformin, 685
  pharmacogenomics and stratified medicines, 685
  platelet reactivity, 518
  and TB, 843
Diastolic BP (DBP), 747
Diazepam, 7–9, 818, 825, 827, 1000
Dibucaine, 822
Dicationic molecules, 913
Diclofenac, 353–354, 445–449
Difference gel electrophoresis (DIGE) method, 153–155
Differential expression, 135
Differentially methylated regions (DMRs), 216
DiGeorge syndrome critical region gene 8 (*DG*CR8), 47–48
DIHS/DRESS/HSS phenotype, 440–441, 452–453
Dihydrofolate reductase (DHFR), 655
Dihydroorotate dehydrogenase (DHODH), 661
Dihydropyridine-sensitive L-type voltage-dependent calcium channel (DHPR), 414
Dihydropyrimidine dehydrogenase (DPD), 355–356
1,25-Dihydroxyvitamin D3 (calcitriol), 25, 394
4,4-Dimethyl-4-silapentane-1-sulfonic acid (DSS), 4, 197
Direct thrombin inhibitors, 487–488, 491–493
  argatroban, 488
  bivalirudin, 488
  dabigatran, monitoring the anticoagulant effect of, 492–493
  interactions, 491–492
  nonhaemorrhagic side effects, 493
  pharmacokinetics, 491
  pharmacology, 491
Direct-acting antiviral therapy, for genotype 1 HCV, 869–871
Directly Observed Treatment Short-Course (DOTS), 835
Direct-to-consumer (DTC) testing, 921
Disability-adjusted life years (DALYs), 618
Discoid lupus erythematosus (DLE), 667
Discovery and validation, 167–172
  cancer markers, 167
  cardiovascular diseases, 167–170
  kidney diseases, 170–171
  neurological complexes, 171–172
  preeclampsia, 171
Disease activity score (DAS), 668–670
Disease modeling, use of stem cells for, 219–221
Disease-associated gene polymorphisms, 705–706
Disease-modifying antirheumatic drugs (DMARDs), 654
  biologic, 668–674
    rituximab, 673–674
    tocilizumab, 674
    tumor necrosis factor antagonists, 668–673
  nonbiologic, 655–668
    azathioprine (AZA), 663–665
    hydroxychloroquine (HCQ), 666–668
    leflunomide, 661–663
    methotrexate (MTX), 655–661
    sulfasalazine (SSZ), 665–666
Distribution, 342–343
  plasma-protein binding, 342–343
  volume of distribution (Vd), 343
Diversifying selection theory, 979
Dizygotic (DZ) twins, 407
DNA damage repair and replication (DDR), 809
DNA library preparation, 127
DNA methylation, 57–61, 99, 136–137
  DNA methyltransferases, 60
  effects of, 58–60
  techniques for measuring, 60–61
DNA-damaging agents, 809–810
DNMT2, 60
Docosahexenoic acid (DHA), 741
Donepezil, 581
Dopamine receptor D4 gene (DRD4), 99
Dopamine receptors, 548
Dopamine transporter (DAT), 549
Dosage of drug, predicting, 10
“Dose revision algorithm”, 503
Dose–response relationship, 370–371
DOTS-Plus, 842
Doxepin, 7–9
Doxorubicin HCl, 357–358
D-penicillamine, 445–449
DpnI enzyme, 64
DraI polymorphism, 851
Drospirenone and Ethinyl Estradiol, 7–9
Drug development, 309
Drug hypersensitivity, stratified approaches to prevent, 455–458
  criteria for stratified approaches, 455–457
  current recommendations for HLA screening, 457–458
Drug hypersensitivity syndromes (DHS), 440–442
  DIHS/DRESS/HSS, 440–441
  immunopathogenesis of, 442–444
    altered peptide repertoire model, 444
    hapten/prohapten model, 442–443
    p-i model, 443–444
  single-organ, 441–442
  SJS/TEN, 441
Drug induced liver injury (DILI), 421–424, 441–442, 453–454, 467
  amoxicillin clavulanate (AC), 422–423
  -associated genetic variants, 471–472

    - DILI genetics in candidate gene era, 471–472
    - functional characterization of DILI-associated alleles, 472
    - human leukocyte antigen genes and DILI risk, 472
  - difficulties in identifying biomarkers for, 470–471
    - acquiring DILI cases, 470–471
    - diagnosis and phenotyping, 471
    - DILI heterogeneity, 471
    - lack of adequate model systems, 471
  - epidemiology and public health impact, 467–468
  - etiology, 468–469
  - flucloxacillin, 423
  - future directions, 473–475
    - epigenomics, 474
    - expanded-content genotyping arrays, 473–474
    - induced pluripotent stem cells (iPSCs), 474–475
    - next-generation sequencing, 474
  - genetic susceptibility to, 850
  - lumiracoxib, 423–424
  - mechanisms for, 468
  - potential value
    - of DILI diagnostics, 470
    - of predictive biomarkers, 469–470
    - of understanding causes of DILI, 469
  - predicting, 853
  - utility of DILI risk variants in the clinic, 472–473
  - ximelagatran, 423
- Drug Induced Liver Injury Network (DILIN), 470–471
- Drug interactions, 731
- Drug metabolism-related genes, 576–577
- Drug safety, 3–4
- Drug transport polymorphisms, 829
- Drug transporters, 357–360, 580
  - ABC transporters, 357–359
    - ABCB1, 357–358
    - ABCC1 and ABCC2, 358
    - ABCG2, 358–359
  - SLC transporters, 359–360
    - organic anion transporters (OATs), 360
    - organic anion–transporting polypeptides (OATPs), 359
    - organic cation transporters (OCTs), 359–360
- Drug-eluting stents (DES), 511, 515–516
- Drug-induced hemolytic anemia, 377
- Drug-induced LQTS (diLQTS), 787, 791
  - congenital LQTS genes in, 795–796
  - prevention, monitoring and treatment of, 795
- Drug-induced skin injury, 417–421
  - abacavir hypersensitivity, 417–419
  - allopurinol hypersensitivity, 420
  - antiepileptic drugs (AEDs), 421
  - carbamazepine hypersensitivity, 419–420
  - maculopapular exanthema reactions, 420–421
    - antituberculosis medications, 420–421
- Drug-induced torsades de pointes (diTdP), 787
- Drug-metabolizing enzymes (DMEs), 937, 985–987, 999–1002
  - in China
    - CYP2C19, 999–1001
    - CYP2D6, 1001
    - CYP3A, 1001
    - histamine N-methyltransferase (HNMT), 1001–1002
- Drug-metabolizing genes, 1039
  - catechol-O-methyltrasferase (COMT), 1049
  - CYP1A1 and CYP1A2, 1039
  - CYP2A6, 1045
  - CYP2C19, 1046
  - CYP2C8, 1045
  - CYP2C9, 1046
  - CYP2D6, 1046
  - CYP2E1, 1047
  - CYP3A4 and CYP3A5, 1055
  - GSTM1, GSTT1, and GSTP1, 1048
  - N-acetyl transferase enzymes (NAT1/NAT2), 1049
  - sulphotransferase 1A1 (SULT1A1), 1049
  - thiopurine S-methyltransferse (TPMT), 1049
  - UGT1A7, 1049
  - uridine diphosphate glucuronosyl transferase 1A1 (UGT1A1), 1048
- Drug-metabolizing P450s, regulation of, 389–391
  - CYP2C8, 389
  - CYP2E1, 389–390
  - CYP3A4, 390–391
- Drug-resistant TB, 841, 844–845
- Drugs associated with QT prolongation, 792–793
- Drugs for neglected diseases initiative (DNDi), 913
- Drug-target genes, 1051–1052
  - angiotensin-converting enzyme (ACE), 1051–1052
  - *CYP19*A1, 1051
  - methylene tetrahydrafolate reductase (MTHFR), 1051
  - *VKOR*C1, 1051
- Drug-transporter genes, 1050
  - *ABC*B1 gene, 1050
  - *SLC22*A1, 1050
  - SLCO1B1, 1050
- D-sotalol, 796
- Dual-specificity phosphatase 1 (*DU*SP1) gene, 644
- Dutch Rheumatoid Arthritis Monitoring (DREAM), 672–673
- Dynamic programming, 262
- Dystrobrevin binding protein 1 (*DTNB*P1 gene), 551

E

- Early after depolarization (EAD) oscillations, 789
- EB101 vaccine immunization, 573
- EBs (embryoid bodies), 221
- Ecarin clotting time (ECT) assays, 492
- Echinocandins, 884–885
- Economic challenges, 19–21
- Efavirenz (EFV) pharmacogenetics, 989–991
- Efavirenz, 353, 445–449
- Efficacy, 372
  - predicting, 5–10
- Eflornithine, 912
  - and nifurtimox, 912–913
- EGFR gene, 376
- Eicosapentenoic acid (EPA), 728
- Electrocardiogram (ECG), 779–780
- Electron capture dissociation (ECD), 148
- Electron impact (EI), 196
- Electron impact ionization, 196
- Electron transfer dissociation (ETD), 148
- Electrospray ionization (ESI), 153, 155–158, 196
- Elongation step, 45
- Eltrombopag, 7–9
- Embryonic stem (ES) cells, 207, 209
- *EML4*–ALK gene, 376
- Empirical Medicine, 14–15
- ENCODE Project Consortium, 38, 42
- Encompass MDx™ platform, 534
- Encyclopedia of DNA Elements (ENCODE) project, 54–56
- “Endocardial abnormalities”, 483–484
- Endomembrane system, 30
- Endoplasmic reticulum (ER), 31, 438
- Endothelin-1 (*E*DN1), 754
- “Enhanced diagnostic” approach, 887
- Enhancer of zeste homolog 2 (EZH2), 213
- Enzyme induction, 346
- Enzyme inhibition, 346
- Enzymes, 369
- EPH receptor B1 (*EP*HB1), 554
- Epidermal growth factor receptor (EGFR) signaling, 316
- Epigenetic disorders and therapies, 65–66
- Epigenetic variation among pluripotent stem cells, 215–216
- Epigenetics, 57, 972–973, 992–993
  - DNA methylation, 57–61
    - DNA methyltransferases, 60
    - effects of, 58–60
    - techniques for measuring methylation, 60–61
  - epigenetic disorders and therapies, 65–66
  - histone PTMs, 63–64
    - effects, 64
  - histones, techniques for measuring, 64–65
- Epigenomic applications, 135–137
  - ChIP-seq, 136
  - DNA methylation, 136–137
- Epigenomics, 129
  - in pluripotent stem cells for regenerative medicine, 216–218
- Epinephrine, 753
- Epipodophyllotoxin (etoposide, VP-16), 357–358
- Epirubicin, 12–13
- Epistatic eQTL interactions, 119
- Epithelial sodium channel (ENaC), 759
- Epoxyeicosatrienoic acid (EET), 393
- Epstein-Barr virus (EBV), 112
  - producer cell lines, 92
  - transformation effects on B cells, 95–96
    - gene expression changes, 96
    - immunophenotypes, 95–96
  - -mediated B-cell transformation, 90–92

    - EBV producer cell lines, 92
    - LCL generation protocols, 90
    - transformation efficiency, 90–92
- EQTLs, identification of, 81–82
- Erlotinib, 7–9, 18
  - responsiveness, 810
- Erythromycin, 794
- eSensor 2C19 Genotyping Test, 534
- Esmolol, 819
- Esomeprazole, 7–9
- 17β-Estradiol, 5–10
- Estrogen Receptor Polymorphisms, 662
- Etanercept (ETN), 668
- Ethambutol, 420
- Ethanol, 389
- Ethics of biomedical research, 934
- Ethnic fractionalization (EF), 974–976
- Ethnicity, 973
- Euchromatin, 43
- Eukaryotes, 29
- Eukaryotic cells, 30
  - membrane-bound organelles in, 30
  - nonmembranous organelles in, 30
- Eukaryotic Promoter Database (EPD), 278
- European Malignant Hyperthermia group (EMHG), 414–415
- European Medicines Agency (EMA), 18–19, 953–954, 957, 959
- European Nucleotide Archive (ENA), 277
- European Science Foundation (ESF), 922–923
- European Zebrafish Resource Centre, 79
- Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group, 924
- Everolimus, 7–9
- Excitation-contraction coupling (ECC), 378–379, 414, 822
- Excretion, 346–348
  - biliary excretion, 348
  - renal excretion, 346–348
- Exemestane, 7–9
- Exercise-induced heart rate, 1006
- Explanatory RCTs (ERCTs), 316
- Exponentially modified protein abundance index (emPAI), 162
- Exportin-5, 48, 385
- Exposure, 847
- Expressed sequence tag (EST), 148
- Expression phenotype studies, 99
- Expression quantitative trait loci (eQTL), 111–112, 417
- Extensively drug-resistant (XDR TB), 842
- Extrapulmonary TB (EPTB), 839–840
- Ezetimibe, 734–736

F

- Factor Xa inhibitors, 493–494
  - bleeding and reversal of anticoagulant effect, 494
  - interactions, 493
  - mechanism of action, 493
  - monitoring of anticoagulant effect, 493–494
  - nonhemorrhagic adverse effects, 494
  - pharmacokinetics, 493
- Facultative heterochromatin, 43
- Familial chylomicronemia syndrome (FCS), 726
- Familial defective apo B-100, 724–725
- Familial dysautonomia, 220
- Familial dysbetalipoproteinemia, 726–727
- Familial hyperaldosteronism type 1 (FH-I), 749–752
- Familial hypercholesterolemia (FH), 722–724
- Familial hypobetalipoproteinemia (FHBL), 725
- Familial LCAT deficiency, 726
- Family-wise error rate (FWER), 253
- FastQ format, 129, 282
- FDA Adverse Event Reporting System (FAERS), 802
- FDA Modernization Act (FDAMA), 938, 944
- FDG-PET, 569
- Fearon’s method, 974–976
- Federal regulations and national healthcare priorities, 943–945
- Fentanyl, 816
- Feprazone, 445–449
- Fetal bovine serum (FBS), 211
- Fetal stem cells, 208–209
  - amniotic fluid (AF) stem cells, 209
  - amniotic membrane stem cells, 209
  - embryonic stem (ES) cells, 209
  - umbilical cord blood (UCB) stem cells, 209
  - Wharton’s jelly stem cells, 209
- Fexinidazole, 913
- Fexofenadine, 799
- Fibrates, 741
- Fibrinogen-beta (FGB) gene, 766–768
- Fibrinolysis, 482
- Filter-aided sample preparation (FASP) method, 151–152
- Fine mapping and resequencing, 255–256
- First-generation sequencing, 125–126
- Fisheye Disease, 726
- FK506 (Tacrolimus), 91
- “Flow abnormalities”, 483–484
- Flucloxacillin, 12–13, 422–423, 441–442, 445–449, 457–458
- Fluconazole, 882, 891, 893, 895
- Flucytosine, 885
- Fluorescence-activated cell sorting (FACS), 218
- Fluoroquinolones, 794
- Fluorouracil, 7–9
- 5-Fluorouracil (5-FU), 376, 807–809
- Fluoxetine, 7–9
  - and Olanzapine, 7–9
- Flurbiprofen, 7–9
- Fluvastatin, 353–354
- Fluvoxamine, 7–9, 333
- Flux analysis, 187
- Fold recognition, 275
- FOLFIRINOX, 807–809
- Folylpolyglutamyl synthase (FPGS), 655
- Fondaparinux, 487
  - adverse effects, 487
  - dosing and reversal, 487
  - pharmacokinetics, 487
  - structure and mechanism of action, 487
- Food and Drug Administration (FDA), 955–957, 959
- Fosinopril, 755–756
- FRAPPE analysis, 980–981
- Free induction decay (FID), 187
- Frequency-dependent selection theory, 979
- Fresh frozen plasma (FFP), 493
- F-statistics (FST), 981–982
- Fulvestrant, 7–9
- Functional eSNPs, identification of, 118–119
- Fungal allergens and asthma, 621
- Fungal diseases
  - public health impact of, 886
  - stratified medicine for, 886–887
    - current successes in, 895–896
- *Fusar*ium, 883

G

- G protein coupled receptor families, 370
- G protein families and their effects, 369
- G protein target, 370
- G protein-activated inwardly rectifying K^+ (GIRK), 1032
- G protein-coupled receptor (GPCR), 368, 739, 761–762
  - kinases, 374, 761–762
- Galantamine, 7–9, 581
- Gambiense HAT, 906
- Gametes, 29–30
- Gamma glutamyl hydrolase (GGH), 655
- Gamma-aminobutyric acid (GABA) receptor genes, 1008–1009
- Gas-chromatography mass spectrometry (GC-MS), 183, 196–197
  - limitations of, 197
- Gasping syndrome, 936
- GATK software, 132
- Gefitinib, 7–9
- Gel-based approaches, 152–155
  - difference gel electrophoresis (DIGE) method, 153–155
  - one-dimensional, 152–154
  - two-dimensional polyacrylamide electrophoresis (2D-PAGE), 154
- Gemcitabine, 12–13, 807–809
- Gemfibrozil, 708
- GenBank, 277
- Gene expression, 45–46
  - transcription, 45
  - translation, 45–46
- Gene Expression Omnibus (GEO), 116–117
- Gene expression QTL (eQTL) mapping, 81
- Gene ontology (GO), 280
- Gene panels, 771
- Gene regulation, 46–51
  - chromatin modifications and gene expression, 50–51
  - by miRNAs, 47–48
  - mRNA targets, 48–50
    - mechanisms of silencing, 49–50
  - nucleosome positioning and occupancy, 51–56
- Gene structures, 44–45
  - number of genes, 44–45
- Genetic basis of drug response and adverse effects, 5

Genetic diversity, 979
Genetic drift, 979
Genetic Information Nondiscrimination Act (GINA), 946–947
Genetic risk prediction score (GRPS), 82
Genetic risk scores, 768–769
Genetic syndromes, 784–787
  congenital long-QT syndrome (cLQTS), 784–787
Genetic variations, types of, 240
Genetic/genomic resources, 79–80
Genetics, basic, 29
  cell, overview of, 29–30
  cell reproduction: cell cycle and mitosis, 30–33
  eukaryotic cells
    membrane-bound organelles in, 30
    nonmembranous organelles in, 30
  meiosis, 33–35
    genetic variation introduced by, 34–35
  Mendelian laws, 35–38
    law of independent assortment, 37
    law of segregation, 35–37
    Mendelian patterns of inheritance in humans, 37–38
  nucleus and chromosomes, 29–30
  public databases for biomedical research in humans, 38–39
Genome sequencing, 812–813
Genome-wide association studies (GWAS), 12, 81–82, 112, 115, 249, 420, 439, 471, 672–673, 683–684, 722, 800–801, 859, 904–905, 1006
  antidepressant drugs (AD), 544–545, 553–554
  in asthma, 628–629, 646, 985
  in pharmacogenomics, 12–13
Genomic applications, 131–133
  *de n*ovo genome sequencing, 131
  variant calling applications, 131–133
    example, 132–133
Genomic variation, 51–56
  copy number variants (CNVs), 54
    mechanisms of formation of, 54
    population genetics of, 54
    and transcriptional regulation, 54
  Encyclopedia of DNA Elements (ENCODE) project, 54–56
  single-nucleotide polymorphisms (SNPs), 53–54
    on gene expression, 53
    linkage and haplotype blocks, 53–54
Genotype testing, 415
Genotype-Tissue Expression (GTEx) project, 116–117
Germline-derived pluripotent stem (gPS) cells, 210
Glibenclamide, 701, 1055
Glipizide, 701
Global Initiative for Asthma (GINA) guidelines, 619–620
Glomerular filtration, 346
Glomerular filtration rate (GFR), 346
Glomerulonephritis, 170–171
Glossina, 902
Glucocorticoid-remediable aldosteronism, 749–752
Glucocorticoids (GCs), 12–13, 83
Glucose-6-phosphate dehydrogenase (G6PD), 5, 905, 377–378
Glutamatergic system, 549–550
Glutathione S-transferase theta (GSTT1), 113–115
Glutathione S-transferases (GSTs), 345–346, 356–357, 472, 664, 819–820, 1048
Glycerol, 389
Glycine amidinotransferase (*G*ATM), 417
Glycogen synthase kinase 3 alpha (*GS*K3A), 552
Glycolysis, 195
*G*NB3, 759
Gold sodium thiomalate, 445–449
Golgi apparatus, 31
Golimumab (GOL), 668
Golm Metabolome Database (GMD), 196–197
*GPI*IIa gene, 486
G-protein coupled receptor kinase 2 (*G*RK2), 639
Grapefruit juice, 817
Graphical user interface (GUI), 269
Grepafloxacin, 794
*G*RK4, 761–762
Gross domestic product (GDP), 19
GSNAP, 134
GSTM1, 1048
*GS*TM1 and *GS*TT1 polymorphisms, 473, 851–852
GSTP1, 1048
*GS*TT1 gene, 113–115, 1048
Guanosine diphosphate (GDP), 46
Guanosine triphosphate (GTP), 46
GWA studies, reporting of, 256
GWAS signals, 760–761

H

H1 linker histone, 62
H2A Barr body-deficient (H2A-Bbd) variant, 62
H3K27 methylation, 213
H3K27me3, 213
H4K16ac levels, 65–66
Half-life, 348–349
Halothane-induced hepatitis, 827
HapMap collections, 94
HapMap project, 242, 859
Hapten/prohapten model, 442–443
Haptoglobin-binding receptor (HPR), 903
Hardy, Godfrey H., 242–243
Hardy-Weinberg equilibrium (HWE), 242–243, 245–246, 326
  in association studies, 243
HCA2, 739
HCV-HIV coinfection, 868
HDL-cholesterol, 721–722, 727–728, 738–741
  apolipoproteins, 721–722
  triglycerides, 721
Hemostasis, 479–482
  "cascade" model of, 480–481
  contemporary view of, 481
    amplification, 481
    initiation, 481
    propagation, 481
  fibrinolysis, 482
  regulation of coagulation, 481–482
Heparin. *See also* Unfractionated heparin (UFH), 936
Heparin-induced thrombocytopenia (HIT), 485–486
Hepatic lipase deficiency, 727
Hepatic metabolic enzymes, 224
Hepatitis C virus (HCV), 859–860
  clinical translation, 866–872
    direct-acting antiviral therapy, 869–871
    genomics and drug development, 871–872
    of *I*TPA genotyping, 871
    peginterferon and ribavirin therapy, 866–869
  genetic determinants of liver fibrosis progression, 865–866
  *IL*28B polymorphism and, 860–864
    clinical phenotypes, 863
    genotype 1 HCV treatment, 860–862
    mechanistic studies and *IFN-lambda*-4, 863–864
    spontaneous HCV clearance, 862–863
  interferon-free therapy for HCV-1, 869–871
  *I*TPA polymorphisms and ribavirin-induced hemolytic anemia, 864–865
Hepatocyte growth factor (HGF), 223
Hepatocytes, 218
Hepatotoxicity, 730
  isoniazid-induced, 849–850
  use of stem cells in studies of, 223–224
*H*ER2 amplification, 810
HERG channel mutations, functional deficits caused by, 799
HERG channels, in clinical application of new drugs, 799–800
HERG potassium channel, 798–799
Heterochromatin, 43, 62
Heterogeneity of metabolites, 195
Heterogeneous stocks, 75
Heteronuclear single-quantum coherence (HSQC), 190
Hexane, 389
H-FABP, 909
High on-treatment platelet reactivity (HTPR), 924–925
High-affinity pentamidine transporter (HAPT1), 910–911
High-energy collision-induced dissociation (HCD), 160
Highly active antiretroviral therapy (HAART), 843, 989
High-platelet reactivity (HPR), 511
Histamine N-methyltransferase (HNMT), 1001–1002
Histone acetyltransferase (HAT), 63–64
Histone demethylase (HDM) activity, 66
Histone methyltransferases (HMT), 63–64
Histone PTMs, 63–64
Histones, techniques for measuring, 64–65
HIV protease inhibitors, 1025
HIV/AIDS

  impact on tuberculosis, 843, 849
HLA-B gene, 937
HLA-B*15:02, 451
*HLA-B*1*502 allele, 419
HLA-B*57:01 molecule, 450
HLA-B*57:01 screening, 457
*HLA-B*5*701, 473
HLA-DRB1, 830
*HLA-DRB1*0*701, 473
*HLA-DRB1*1*501, 473
HMG-CoA reductase inhibitors, 731
*HM*GCR gene, 102
Homology modeling, 272–275
  alignment correction, 273
  backbone generation, 273
  loop modeling, 273–274
  model optimization, 275
  model validation, 275
  side-chain modeling, 274–275
  template recognition and initial alignment, 273
Homozygosity, 75
5HTT function, 1032
Human African trypanosomiasis (HAT), 901–903
  diagnosis of, 907–908
  future perspectives, 913–914
  genetics of trypanosome lytic factors and parasite factors, 903–905
  immune responses and immunoavoidance, 905–906
  pathologies associated with, 906–907
  staging the disease, 908–909
  treatment, 909–913
    compounds in clinical or advanced preclinical trials, 913
    current drugs, 909–910
    stage 1, 910–911
    stage 2, 911–913
Human dihydrofolate reductase (DHFR) gene, 396
*Human ether-à-go-go g*ene channel (HERG), 784, 788
Human genetic variation, 979
Human genome, structure of, 41–44
  chromatin structure, 42–43
  nucleosome, 43–44
Human Genome Project, 859, 931
Human hERG structure and polymorphisms., 1009
Human immunodeficiency virus (HIV) infection, 417
Human leukocyte antigen (HLA), 92–93, 417, 852
  and hypersensitivity reactions, 444–455
    abacavir, 444–450
    allopurinol, 452–453
    amoxicillin-clavulanate, 453–455
    carbamazepine, 450–452
    nevirapine (NVP), 453
  diversity, linkage disequilibrium, and specificity of peptide binding, 439–440
Human Metabolome Database, 279
Human rhinovirus (HRV), 622
Human sex chromosomes, 33
Human umbilical cord blood neural stem cell line (HUCB-NSC), 225
Huntington's disease, 65
HyBeacons® Assay, 534
Hybridization capture, 139
Hydralazine, 445–449
Hydrochlorothiazide (HCTZ), 12–13, 763, 755–756
  blood pressure response to, 760–761
Hydrolysis, 345
3-Hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA reductase) inhibitors, 415–416
Hydroxychloroquine (HCQ), 666–668
  CYP, 667
  metabolism, 667
  pharmacogenetics of, 668
  TNF/IL-10, 667–668
5-Hydroxymethylcytosine (5hmC), 60
Hydroxymethylglutaryl-coenzyme A (HMG-CoA) reductase, 728
5-(4'-Hydroxyphenyl)-5-phenylhydantoin (p-HPPH), 410
Hypersensitivity syndrome (HSS), 417
Hypertension, 747
  antihypertensive treatment, 764–771
    clinical trials, 764
    genes with strong evidence of association with, 764–771
  blood pressure response to antihypertensive drugs, 758–764
    β-blockers, 761–762
    ACE inhibitors and ARBs, 763–764
    calcium channel blockers (CCB), 762–763
    thiazide diuretics, 758–761
  definition, 747–748
  epidemiology and public health impact, 747–749
    heritability of blood pressure, 748–749
    hypertension causation, 748
    risk factors, 748
  future perspectives, 771–772
  genetic basis of, 749–754
    autonomic nervous system, 753
    sodium and vascular volume, 749–753
    vascular mechanisms, 753–754
  stratified approaches, need for, 754–758
    age and race, 756–757
    genetic approaches, 758
    nongenetic approaches, 755–758
    plasma renin activity, 756
    pretreatment blood pressure, 755–756
Hypoglycemia, 704
Hypothalamic−pituitary−adrenal (HPA) axis and inflammation, 550–551
Hypoxanthine phosphoribosyl transferase (HPRT), 663–664
Hypoxia inducible factor 1α (HIF-1α), 392
Icahn School of Medicine at Mount Sinai, 927

I

ICH E16, 951–953
I-Densy™ IS-5320, 534
Identical by descent (IBD), 246–247
Identifiability, 75
IGF2r gene, 59
IKBKAP gene, 220
I_{Kr} blockers, 790
I_{Kr}-blocking drugs, 788
*IL*28B genotyping, 866
  and first-generation protease inhibitors, 869
  and next-generation DAA response, 869–871
  and pharmaco-economics, 871
  and treatment response, 870
*IL*28B polymorphism and HCV outcomes, 860–864
  clinical phenotypes, 863
  genotype 1 HCV treatment, 860–862
  mechanistic studies and *IFN-lambda*-4, 863–864
  spontaneous HCV clearance, 862–863
Illumina, 127–128
Illumina HiSeq platform, 132
Iloperidone, 7–9, 12–13
Imatinib, 7–9, 16–17
Imatinib mesylate, 18
Imipramine, 7–9
Immunosuppressants, 1055–1056
Immunotherapy, source of antigen-presenting cells for, 103
Implantable cardioverter-defibrillator (ICD) insertion, 787
Imprinting centers (ICs), 215
Imputation, 254–255
In vivo amphotericin B, 881
In vivo models, in pharmacogenomic research, 81–82
  identification of eQTLs, 81–82
  quantitative trait mapping, 81
Inbred strains, 75, 81
Indacaterol, 7–9
India, pharmacogenomics in, 1037–1038
  future directions, 1056–1057
  linguistic distribution of Indian populations., 1039
  pharmacogenetic studies, 1052–1056
    analgesics, 1056
    anticancer drugs, 1052–1053
    antifolates, 1056
    cardiovascular drugs, 1053–1054
    immunosuppressants, 1055–1056
    oral hypoglycemic drugs, 1054–1055
  population structure and diversity, 1038–1039
    drug-metabolizing genes, 1039
    drug-target genes, 1051–1052
    drug-transporter genes, 1050
    frequency distribution and pattern of genetic variants of pharmacogenes, 1039–1045
Indinavir, 354
Individuality, 75
Individualized Molecular Pancreatic Cancer Therapy (IMPaCT), 810–811
Individualized/personalized medicine, 14–15
Induced pluripotent stem cells (iPSCs), 474–475, 210–211
  source material of, 103–104
INFINITI CYP2C19 assay, 534

Inflammation, platelet hyper-reactivity, 518
Inflammatory cytokines, 551
Infliximab (INF), 668
Informed consent, 938–940
Inhalational anesthetic agents, 827–828
Inhaled corticosteroids (ICS), 629–630
Inheritance in humans, Mendelian patterns of, 37–38
Initiation of transcription, 45
Inosine triphosphate (ITP), 663
Institutional review boards (IRBs), 940–941
Insulin, 683
Integrated Genomics Viewer (IGV), 276
Intention-to-treat (ITT) analysis, 992
Interferon a (Pegasys), 12–13
Interferon a2b/ribavirin, 12–13
Interferon-free therapy for HCV-1, 869
*Interferon-lambda-4 (IF*NL4)
  *IL*28B and, 863–864
Interindividual variation of drug response, 4
Intermediate filaments, 30
Internal standard (IS), 197
International Cell Line Authentication Committee (ICLAC), 93
International Clopidogrel Pharmacogenomics Consortium (ICPC), 301
International Consortium for Antihypertensives Pharmacogenomics Studies (ICAPS), 301
International HapMap Project, 38, 931
International Histocompatibility Working Group (IHWG), 95
International Mouse Strain Resources (IMSR), 79
International Normalized Ratio (INR), 413, 485, 490
International Nucleotide Sequence Database Collaboration (INSDC), 276
International Sensitivity Index (ISI), 490
International Union of Basic and Clinical Pharmacology (IUPHAR), 366
International Warfarin Pharmacogenetics Consortium (IWPC), 10, 301, 413, 503
Interphase, 33
Intra-family correlation coefficient (ICC), 237
Intravenous immunoglobulins (IVIgs), 106
Invader® Assay, 534
Invader® UGT1A1 Molecular Assay, 944–945
Invasive fungal disease (IFD), 886–887
Inward rectifier K+ *channel Kir*6.2, 705
Ion channels, 368–369
  common variants and common polymorphisms in, 800
Ion mobility spectrometry (IMS), 195
Ion Personal Genome Machine (PGM), 126–127
Ion Proton systems, 126–127
Ion Torrent, 126–127
Ionization techniques, 157–158
  electrospray ionization (ESI), 153, 157–158
  matrix-assisted laser desorption ionization (MALDI), 153, 158
iPath, 200
IPS cells, 215
  -derived cardiomyocytes, 222
Irinotecan, 7–9, 412–413, 809
Iscove's Modified Dulbecco's Medium (IMDM), 211
ISG expression, 864
Isogenicity, 75
Isomers, 195
Isoniazid (INH), 7–10, 389, 411–412, 420, 441–442, 445–449, 845–852
  acetylator genotypes and phenotypes, 846–847
  *CYP*2E1 polymorphisms, 851
  genetic susceptibility to drug-induced liver injury, 850
  GSTM*1*, GSTT1 polymorphisms, 851–852
  human leucocyte antigen alleles, 852
  -induced hepatotoxicity, 849
    mechanism of, 849–850
  isoniazid acetylator status
    in pediatric TB, 849
    and treatment efficacy, 847–849
  manganese superoxide dismutase (MnSOD), 852
  metabolism, 846
  *N*AT2 polymorphisms, 850–851
  tumor necrosis factor (TNF) polymorphisms, 852
  UDP polymorphisms, 852
Isoniazid drug-induced liver injury (INH-DILI), 992
Isosorbide and Hydralazine, 7–9
Isotope-coded affinity tag (ICAT) method, 159–160
*I*TPA alleles, 664
*I*TPA genotyping, clinical translation of, 871
*I*TPA polymorphisms
  and ribavirin-induced hemolytic anemia, 864–865
Itraconazole, 882–883, 891, 894–895
iTRAQ, 153, 160
Ivacaftor, 7–9

J

JAK2 gene, 376
JAX mice database, 79
J-coupling, 190
Jenner, Edward, 933
J-resolved NMR, 190

K

*KC*NH2 gene, 784
*KC*NH2 polymorphisms, 798–799
  and disease susceptibility, 796
  drug sensitivity-related, 799–800
KCNH2 potassium channels, 1009–1010
KCNJ11 polymorphism, 1010
*KCN*MB1 genotype, 762–763, 769–770
*KC*NQ1 gene, 784, 1010
Ketoconazole, 817
Kidney diseases, 170–171
Kinesin-like protein 6 (KIF6), 733
King, Helen Dean, 74
Kir6.2, 1010
*K*RAS gene, 395–396

L

Label-free methods, 153, 161
  design of MRM transitions/MRM assays, 162
  generation of SRM method, 162
  label-free using ion intensities, 162–163
  label-free using spectral counting, 163
Lactase persistence/nonpersistence, 984
Lamotrigine (LTG), 418, 421, 445–449
Lanosterol, 716
Lansoprazole, 7–9, 329–330
Lapatinib, 7–9, 422, 424, 445–449
Latin American countries, 1015
  PGx studies in, 1015–1016
  Population, area, economic facts and scientific output of, 1016
Law of independent assortment, 37
Law of segregation, 35–37
LCL generation protocols, 90
LCL transcriptomics, 98–99
LC-MS system, 153, 156
LDL-C, 727–731
Lecithin-cholesterol acyltransferase (*L*CAT) gene, 718, 726
Leflunomide, 661–663
  *AB*CG2, 661–662
  Cytochrome P450 System, 662–663
  *DH*ODH, 661
  estrogen receptor polymorphisms, 662
  metabolism of, 661
  pharmacogenetics of, 662
Lenalidomide, 7–9
LEOPARD syndrome, 220, 222
Letrozole, 7–9, 1052–1053
Leukotriene C4 synthase, 642
Leukotriene modifier drugs, 642–643
Leukotriene receptor antagonist (LTRA), 632, 640
Leukotriene synthesis inhibitors (LTSIs), 640
Levamisole, 445–449
Library preparation, 133
LightCycler, 534
Limits of detection (LoD), 196–197
Limits of quantification (LoQ), 196–197
Linguistic diversity, in Africa, 973–974
Linkage disequilibrium (LD), 241–243
  Hardy-Weinberg equilibrium, 242–243, 245–246
  HWE in association studies, 243
Linkage studies, 243–247
  identical by descent (IBD), 246–247
Lipid disorders and lipid levels, 722–741
Lipid metabolism, 716–718
  cholesterol synthesis, 716–717
  dysfunction, 568
  endogenous pathway, 718
  exogenous pathway, 717–718
  lipoproteins, 717
  reverse cholesterol transport, 718
Lipid-lowering therapies, 715–716
  future directions, 741–742
  genetic basis of lipid disorders and lipid levels, 722–741
    common genetic variants and lipid levels, 727–728
    HDL-C, 738–741

        - lipids and pharmacogenomics, 728–738
        - primary dyslipidemias, 722–727
    - HDL-cholesterol, 721–722
    - total cholesterol and low-density lipoprotein cholesterol, 719–721
- LIPO window, 185
- Lipophobic drugs, 342–343
- Lipoprotein lipase (LPL), 717–718
- Lipoproteins, 717, 722
- 5-Lipoxygenase, 640–642
- Liquid chromatography–mass spectrometry, 194–196
    - limitations of liquid spectrometry, 194–196
    - sample preparation for, 194
- Liver expression quantitative trait loci (eQTL), 111
    - correlation of, and CYP enzyme activity, 116
    - future directions, 117–120
        - copy number variants, 119–120
        - epistatic eQTL interactions, 119
        - functional eSNPs, identification of, 118–119
        - liver protein QTL, 120
        - miRNA eQTL, identification of, 118
        - next-generation sequencing, 117–118
        - sQTL, identification of, 118
    - in pharmacogenetic and clinical genome-wide association studies, 112–115
    - reproducibility, 115–116
    - tissue specificity, 116
    - web-based liver eQTL data resources, 116–117
- Liver protein QTL, 120
- Liver X receptor-α (LXRA), 770
- Liver-specific transporter 1 (LST-1), 1005
- LMWM window, 185
- Locally advanced pancreatic cancer, 811
- Lomitapide (BMS-201038), 737
- Long interspersed repeats (LINES), 42
- Long-acting bronchodilator, 629
- Long-acting β2-adrenoceptor agonists (LABA), 374–375
- Long-QT syndrome (LQTS), 780, 784, 787
    - acquired, 787
    - congenital (cLQTS), 784–787
    - drug-induced LQTS (diLQTS), 787, 791
        - congenital LQTS genes in, 795–796
        - pharmacogenomic studies, 798–801
        - prevention, monitoring and treatment of, 795
    - public health impact, 796–797
- Long-term stability, 75
- Long-term subculture effects on LCLs, 96–97
    - gene expression phenotypes, 96
    - growth properties, 96
    - miRNA expression phenotypes, 96–97
- Loop modeling, 273–274
- Loop-mediated isothermal amplification (LAMP), 908
- Lopinavir, 1025
- Losartan, 352
- Loss of heterozygosity (LOH), 97
- Low-affinity IgE receptor gene, 644
- Low-affinity pentamidine transporter (LAPT1), 910–911
- Low-density lipoprotein (LDL) cholesterol, 719–721
- Low-density lipoprotein cholesterol (LDL-C), 415–416
- Low-molecular-weight heparin (LMWH), 481, 486–487
- Low-molecular-weight metabolites (LMWM) window, 185
- LT Ion Torrent, 127
- LT SOLiD platform, 127–129
- L-thyroxine in differentiated thyroid cancer, 1028
- Luminex 100/200 assay, 534
- Lumiracoxib, 12–13, 423–424, 445–449
- Lymphadenopathy, 906
- Lymphoblastoid cell lines (LCLs), 89, 112
    - applications, in pharmacogenomic studies, 100
    - biological characteristics of, 95–97
        - EBV transformation effects on B cells, 95–96
        - long-term subculture effects on LCLs, 96–97
    - Epstein-Barr virus (EBV)-mediated B-cell transformation, 90–92
        - EBV producer cell lines, 92
        - LCL generation protocols, 90
        - transformation efficiency, 90–92
    - future perspectives, 105–107
    - generation, in biobanks, 92–93
        - blood cell banking for LCL generation, 92
        - current biobanks, 92
        - quality control of LCLs, 92–93
    - generation of, 90–95
    - potentials and limitations of, 97–105
        - factors to consider, 104–105
        - genomic studies, 97–100
        - immunological studies, 103
        - pharmacogenomic studies, 100–103
        - stem cell research, 103–104
    - repositories and collections, 93–95
        - CEPH collections, 93–94
        - HapMap collections, 94
        - NBK collections, 94
        - NLGIP collections, 94–95
- Lymphotoxins, 668
- Lysosomes, 31

M

- MacFarlane cascade, 480
- MacroH2A variant, 62
- Macrolides, 794
- MACS (Multicenter AIDS Cohort Study), 95
- Maculopapular exanthema (MPE), 417, 420–421
    - antituberculosis medications, 420–421
- Maintenance methylation, 60
- Major adverse cardiovascular events (MACE), 511
- Major depressive disorder (MDD), 543–545, 548
- Major histocompatibility binding (MHC) tests, 423
- Major histocompatibility complex (MHC), 417, 437–438, 653
- Malignant hyperthermia (MH), 822–825, 378–379
    - and anesthesia, 414–415
        - genotype testing, 415
        - muscle contracture testing, 414–415
- Manganese superoxide dismutase (MnSOD), 852
- MapSplice, 134
- Maraviroc, 7–9
- Mass spectrometry, 192–194
- MassTrix, 200
- MATE1:SLC47A1 gene polymorphisms and metformin response, 702
- Material-enhanced laser desorption/ionization (MELDI), 158
- Matrix-assisted laser desorption-ionization (MALDI), 153, 157–158
- Maximov, Alexander A., 207
- MDR-1 polymorphisms, 816
- Meiosis, 33–35
    - genetic variation introduced by, 34–35
- Melarsoprol, 911–912
- Memantine, 582
- Membrane-bound organelles in eukaryotic cells, 30
- Mendelian laws, 35–38
    - law of independent assortment, 37
    - law of segregation, 35–37
    - Mendelian patterns of inheritance in humans, 37–38
- Mercaptopurine, 7–9
- Mesenchymal stem cells (MSCs), 208
- Met allele carriers, 551
- Meta-analysis, 255
- Metabolism, 343–346
    - enzyme induction, 346
    - enzyme inhibition, 346
    - non-CYP450 enzymes, 345–346
    - phase II reactions, 345
    - phase I reactions, 344–345
- Metabolite fingerprinting, 184–185
    - via NMR spectral binning, 184–185
- Metabolome-wide association studies (MWAS), 184
- Metabolomic data, 197–200
    - annotation/identification, 200
    - contextualization of, 200
    - MS-based metabolomics data, 199–200
    - NMR-based metabolomics data, 197–200
        - binning/bucketing, 198–199
        - data normalization, 199
        - spectral deconvolution and peak picking, 198
- Metabolomics, 181
    - analytical platform, 187–197
        - common 1D NMR pulse programs, 189–191
        - gas chromatography–mass spectrometry, 196–197
        - limitations of NMR, 191
        - liquid chromatography–mass spectrometry, 194–196
        - mass spectrometry, 192–194

    nuclear magnetic resonance, 187–189
  applications, in clinical and biomedical research, 183
  biochemistry of metabolome, 182–183
  experimental design for, 183–184
  metabolome and, 181–182
  metabolomic data, analysis of, 197–200
    annotation/identification, 200
    contextualization of metabolomics data, 200
    MS-based metabolomics data, 199–200
    NMR-based metabolomics data, 197–200
  techniques and technologies, 184–187
    fingerprinting, 184–185
    flux analysis, 187
    metabolite pool analysis, 185
    metabolite profiling, 185
    metabolite profiling using NMR, 185–187
Metacarpophalangeal (MCP), 654
MetaCyc, 279–280
Metagenomic applications, 129, 137–141
  metagenomic sequencing, 138–139
  microbiome profiling, 137–138
Metaphase, 32
Metastatic pancreatic cancer, 808–811
  DNA-damaging agents, 809–810
  erlotinib responsiveness, 810
  gemcitabine, 808–809
  *H*ER2 amplification, 810
  individualized molecular pancreatic cancer therapy trial, 810–811
  nab-paclitaxel, 810
MetExplore, 200
Metformin, 12–13, 692, 700–701, 1054–1055
Methadone, 353, 816, 829–830
Methazolamide, 445–449
Methotrexate (MTX), 655–661, 1056
  *AB*CB1, 658
  *ATIC/CC*ND1, 660–661
  cellular pathway of, 656
  MTHFR/TYMS/DHFR, 658–660
  pharmacogenetics of, 656–657
  transporter genes—RFC1/GGH/ *AB*CB1, 657–658
Methotrexate, 12–13
Methylated DNA sequences (MeDIP), 60–61
Methylation, 58–59
  techniques for measuring, 60–61
Methyl-binding domain (MBD), 61
MethylC sequencing (MethylC-seq), 215
Methyl-CpG–binding protein (MBD1), 66
Methyldopa, 445–449
Methylene tetrahydrafolate reductase (MTHFR), 655, 1051
Methylprednisolone (MPL), 84–85
Metoprolol, 7–9, 351
MG-RAST, 138–139
MHC class I molecule (MHCI), 438–439
MHC class II molecule (MHCII), 439
Mibefradil, 796
Micafungin, 893, 895
Microalbuminuria, 171
Microbiome profiling, 137–138
Microelectrode arrays (MEAs), 222
Microfilaments, 30
Microhematocrit centrifugation technique (mHCT), 908
MicroRNA (miRNAs), 45, 47, 49, 100, 118, 279, 385
  biogenesis, 385–386
  as biomarkers and therapeutic targets, 397
  and drug response, 397
  function, 386
  gene regulation by, 47–48
  modulation of, and its toxicological outcome, 394–395
  pharmacogenetics associated with disease or drug response, 395–397
  regulation of cancer-related P450s, 391–394
    CYP19A1, 393–394
    CYP1A1, 391–392
    CYP1B1, 392
    CYP24A1, 394
    CYP2A, 392–393
    CYP2J2, 393
  regulation of drug-metabolizing P450s, 389–391
    CYP2C8, 389
    CYP2E1, 389–390
    CYP3A4, 390–391
  target identification, 386–389
Microsomal triglyceride transfer protein (MTP), 717, 737
Microtubules, 30
Midazolam, 825, 827
Minimal anticipated biological effect level (MABEL), 310–311
Ministry of Health, Labor and Welfare (MHLW), 954
Minor allele frequency (MAF), 239
MiR-103, 389
MiR-107, 389
MirBase, 279
MiRNA eQTL, identification of, 118
MiRNA expression phenotypes, 96–97
MiRNA recognition element (MRE), 386
Mitochondria, 31
Mitogen-activated protein kinases (MAPKs), 671–672
Mitomycin C (MMC), 219
Mitosis, 30–33, 35
Mitoxantrone-resistant protein (MXR). *See* *AB*CG2
Mivacurium, 822
*MM*P9 and *MM*P12, 770
Modafinil, 7–9
Model optimization, 275
Model validation, 275
“Molecular autopsy”, 786
Molecular evolution and phylogenetics, 264–269
  methods of tree inference, 267–269
  principles of tree inference, 266–267
  tree terminology, 265–266
Molecular genetics evolutionary analysis (MEGA), 269
Molecular signatures of stem cells, 212–215
  alteration of gene expression, 215
  chromatin modifications, 214–215
    polycomb group (PcG) protein complexes, 214
    trithorax-(TrxG-) group protein complexes, 214–215
  epigenetic modifications, 212–214
    DNA methylation, 212–213
    histone modifications, 213
    small interfering RNAs (siRNAs), 213–214
  genomic imprinting, 215
Monoamine oxidase (MAO), 548
Monoaminergic system, 544–549
Monoamines, 818
Monogenic dyslipidemias, 741–742
Monozygotic (MZ) twins, 407
Morphine, 816, 829
Morphine to morphine 6-glucuronide (M6G), 820
Motor molecules, 30
Mouse embryonic fibroblasts (MEFs), 211–212, 226
Mouse Genome Informatics database, 79
Mouse models, 74
Mouse Phenome Database, 79
MRNA targets, 48–50
  mechanisms of silencing, 49–50
MS-based metabolomics data, 199–200
MTHFR, 114
*MTHFR/TYMS/D*HFR, 658–660
Multicenter AIDS Cohort Study (MACS), 97
Multidimensional protein identification technology (MudPIT), 153, 156–157
Multidrug resistance protein 2 (MRP2), 888
Multidrug resistance-associated proteins (MRPs), 655, 1002–1003
Multidrug-resistant TB (MDR-TB), 836, 842, 844–845
Multidrug-resistant transporters (MDRs), 655
Multiple endocrine neoplasia syndrome type 2 (MEN-2), 753
Multiple reaction monitoring (MRM), 153, 161–163, 193–194, 196–197
  using triple-quadrupole mass spectrometer, 194
Multipotent stem cells, 208
Muscle contracture testing, 414–415
Muscle toxicity (myopathy), 729–730
  SLCO1B1, 730
Mycobacterium tuberculosis, 835, 841
Mycophenolic acid (MPA), 7–9, 358
Myocardial action potential, Phase 1, 783

N

N-3 polyunsaturated fatty acids. *See* Omega 3 fatty acids
Nab-paclitaxel, 810
N-acetyl procainamide (NAPA), 791
N-acetyl transferase (NAT2), 846
N-acetyl transferase enzymes (NAT1/NAT2), 819, 1049
N-Acetyl Transferase-2 (NAT-2), 10, 992
N-acetyltransferases (NATs), 356
Narodov Mira, 974
Narrow sense heritability, 237–238

*N*AT2 polymorphisms, 850–851, 665–666, 473
Nateglinide, 690, 696
National Commission for the Protection of Research Subjects of Biomedical and Behavioral Research, 934–935
National Human Genome Research Institute (NHGRI), 38, 54–55, 126
National Institute of Child Health and Human Development (NICHD), 936
"Nature versus nurture" debate, 972
NBK collections, 94
NBRP-Rat Phenome Database, 79
Nearest shrunken centroid (NSC), 163
*NED*D4L, 759–760, 765
Nefazodone, 7–9
Negative predictive value (NPV), 456
Nested case-control study, 319
N-ethyl-N-nitrosourea (ENU), 78, 219
Neurofibrillary tangles (NFT), 567, 571
Neurological complexes, 171–172
Neuromuscular blocking drugs (NMBDs), 828
  depolarizing muscle relaxants (DMRs), 828
  nondepolarizing muscle relaxants (NMDRs), 828
Neuronal apoptosis, in AD, 567–568
Neurotoxicity studies, use of stem cells in, 224–226
Neutral theory of evolution, 979
Neutralizing antibodies for serotherapy, 103
Nevirapine (NVP), 12–13, 418, 422, 445–449, 453
New chemical entities (NCE), 221, 224–225
Newborn screening (NBS), 947
Next-generation sequencing (NGS) technology, 61, 117–118, 125, 425
  application-independent data analysis, 129–131
    assembly through read mapping, 130–131
    base calling, 129
    read assembly, 130
    read quality control, 129–130
    sequence and read count, 131
  concept of, 126–129
  epigenomic applications, 135–137
    ChIP-seq, 136
    DNA methylation, 136–137
  first-generation sequencing, 125–126
  genomic applications, 131–133
    de novo genome sequencing, 131
    variant calling applications, 131–133
  metagenomic applications, 137–141
    metagenomic sequencing, 138–139
    microbiome profiling, 137–138
  second-generation sequencing (SGS), 126–129
    applications, 128–129
    clonal amplification, 127
    DNA library preparation, 127
    sequencing DNA fragments in parallel, 127–128
    SGS sample preparation, 128
    spatial separation of fragments, 127
    third-generation methods (TGS), 128
  transcriptomic applications, 133–135
    differential expression, 135
    library preparation, 133
    quantification of transcript abundance, 135
    sequencing, 133–134
    transcript reconstruction, 134–135
"Next-generation sequencing" (NGS), 126
Nicotinic acid (Niacin), 739–740
Niemann-Pick C1-Like 1 (NPC1L1), 734–736
Nifedipine, 354
Nifurtimox-eflornithine combination therapy (NECT), 905, 909–913
Nilotinib, 7–9
Ninjurin2 (*NI*NJ2), 567
Nitric oxide (NO), 905
Nitric oxide synthase (*N*OS3), 754
Nitric oxide synthase 1 adaptor protein (*NOS*1AP) gene, 801
^{15}N-labeled ammonium persulfate, 160
NLGIP collections, 94–95
N-Methyl-D-Aspartate (NMDA) receptor antagonist, 549, 582
N-methyl-N-nitro-nitrosoguanidine (MNNG), 219
NMR hyphenation, 191
NMR-based metabolomics data, 197–200
  binning/bucketing, 198–199
  data normalization, 199
  spectral deconvolution and peak picking, 198
N-nitrosodimethylamine, 389
N-nitrosopyrrolidine, 389
Nociceptin/orphanin FQ peptide (NOP), 828–829
Nomograms, 811–812
Nonalcoholic fatty liver disease (NAFLD), 865
Nonbiologic DMARDS, pharmacogenetics of, 655–668
  azathioprine (AZA), 663–665
    glutathione S transferases (GSTs), 664
    *I*TPA, 664
    purine pathway enzymes, 664–665
    *T*PMT, 664
  hydroxychloroquine (HCQ), 666–668
    CYP, 667
    TNF/IL-10, 667–668
  leflunomide, 661–663
    *AB*CG2, 661–662
    Cytochrome P450 System, 662–663
    *DH*ODH, 661
    Estrogen Receptor Polymorphisms, 662
  methotrexate (MTX), 655–661
    *AB*CB1, 658
    *ATIC/CC*ND1, 660–661
    *MTHFR/TYMS/D*HFR, 658–660
    transporter genes—RFC1/GGH/*AB*CB1, 657–658
  sulfasalazine (SSZ), 665–666
    *N*AT2, 665–666
Noncommunicable diseases (NCDs), 977–978
Non-CYP enzymes, 818–819
  butrylcholinesterase (BChE), 819
  carboxylesterases (CESs), 818–819
  catechol-O-methyltransferase (COMT), 818
Non-CYP450 enzymes, 345–346, 354–357
  dihydropyrimidine dehydrogenase (DPD), 355–356
  glutathione S-transferases (GSTs), 356–357
  N-acetyltransferases (NATs), 356
  thiopurine S-mthyltransferase (TPMT), 355
  UDP Glucuronosyltransferases (UGTs), 354–355
Nondepolarizing muscle relaxants (NMDRs), 828
Non-gel-based approaches, 155–157
Non-genotype 1 HCV, 867–868
Nonhomologous end joining (NHEJ), 78
Noninherited maternal HLA antigens (NIMA), 653
Nonmembranous organelles in eukaryotic cells, 30
Non-receptor-mediated drug response, 373
Nonrodent models of human disease, 79
Nonspecific cholinesterase. *See* Butyrylcholinesterase (BChE)
Non-ST-elevation myocardial infarction (NSTEMI), 509–510
Nonsteroidal anti-inflammatory drugs (NSAIDs), 654, 830
Nonsulfonylurea insulin secretagogues (glinides), 706–707
  organic Anion–Transporting *Polypeptide 1*B1 (OATP*1B*1, SLCO1B1), 706–707
  pharmacodynamics, 707
  pharmacokinetics, 706
Norepinephrine (NE), 548, 753
Norepinephrine transporter (NET), 548
North American Malignant Hyperthermia Group (NAMHG), 414–415
Nortriptyline, 7–9
*N*OS3, 770
*N*PPA, 770
NSAIDs, 445–449, 1021–1024
Nuclear magnetic resonance (NMR), 187–189
  experimental techniques, 188–189
    preanalytical factors, 188
    sample preparation, 188–189
  instrumentation, 188
  limitations of, 191
    matrix effects, 191–192
    spectral overlap, 191
Nuclear Overhauser Effect SpectroscopY (NOESY), 189
Nuclear Overhauser Effects (NOE), 190
Nuclear receptor, classes of, 369
Nucleic acid, 41–42
Nucleic acids and proteins, analysis of sequences of, 260–264
  comparison and alignment of two sequences, 261–263
  database searching, 264
  dot matrix comparison of two sequences, 263
  multiple sequence alignment, 263–264
  small motifs, 261
Nucleoside reverse transcriptase inhibitor (NRTI), 985
Nucleosome, 43–44, 61–62
  positioning and occupancy, 51–56
  structure of, 44
Nucleus, 31
  and chromosomes, 29–30
Number needed to test (NNT), 456
Nuremberg Code, 933, 938

O

OATP1B1-mediated herb–drug interaction, 1011–1012
OATP1B3, 359
OCLA (Omeprazole Clopidogrel Aspirin), 517
OCT2:SLC22A2 gene polymorphisms and metformin response, 702
O-desmethylmetoprolol, 351
Olanzapine, 12–13, 794
Oligopotent stem cells, 208
Omega 3 fatty acids, 741
Omeprazole, 7–9, 329–330, 353–354, 818
Onchocerca volvulus, 911
Ondansetron, 830
Online Mendelian Inheritance in Man (OMIM), 280
Open reading frames (ORFs), 48
Operational taxonomic units (OTU), 137–138
Opioids, 818, 821, 828–830, 1056
  μ receptor polymorphism, 829
  codeine, 829
  drug transport polymorphisms, 829
  metabolism, 829
  methadone, 829–830
  morphine, 829
  tramadol, 830
Optimization of detection, 191
Oral antidiabetic drugs
  genetic basis of diabetes, 683–684
  genetic variants associated with, 685–709
    biguanides (metformin), 685–703
    nonsulfonylurea insulin secretagogues (glinides), 706–707
    sulfonylureas, 703–706
    thiazolidinediones, 707–709
  management of, 684–685
  pharmacogenomics, 685
  stratified medicines, 685
  stratified oral antidiabetic medicines, 709–710
Oral contraceptive pill (OCP), 377
Oral hypoglycemic drugs, 1054–1055
  anticonvulsants, 1055
  glibenclamide, 1055
  metformin, 1054–1055
  pioglitazone, 1055
Organic anion transporter peptides (OATPs), 852
Organic anion transporters (OATs), 348, 360, 580
Organic anion transporting polypeptide 1B1 (OATP1B1), 1005, 1050
Organic anionic transporters (OATs), 890
Organic anionic transporting polypeptides (OATPs), 359, 708, 731, 890, 892
Organic cation transporters (OCTs), 348, 359–360, 1054
Ornithine decarboxylase (ODC), 912
Orphan drugs, 932
Osler, William, 1037–1038
Osteonectin. *See* Secreted Protein Acid and Rich in Cysteine (SPARC)
Osteoporosis, 485
Outbred rodent stocks, 74–75
Out-of-Africa population origin model, 982–984
Oxaliplatin, 809
Oxcarbazepine (OXC), 418, 421, 445–449
Oxicam, 445–449
Oxidations, 344–345
Oxidative stress, 568

P

P generation, 35
P2 amino-purine transporter, 911
P2X7 receptors, 551
P450 2C9, 10
P450 phenotypes, 408–410
  CYP2C9, 410
    phenytoin, 410
  CYP2D6, 408
    β-blockers, 408–410
    codeine, 409–410
  tardive dyskinesia and antipsychotics, 410
PacBio RS system, 128
PAH, pharmacogenomics and, 84
Pain and analgesia, pharmacogenomics of, 83–84
Pamidronate, zoledronic acid, 12–13
Pancreatic cancer (PC), 807
  biorepository and privacy issues, 813
  central knowledge database, 812–813
  current treatment paradigm, 807–808
  preclinical models of, 812
  stratified therapeutic approach, 808–812
    locally advanced pancreatic cancer, 811
    metastatic pancreatic cancer, 808–811
    resectable pancreatic cancer, 811–812
Pancreatic ductal adenocarcinoma (PDAC), 132–133
Panitumumab, 7–9, 18
Pantoprazole, 7–9
Pantothenate kinase (PANK) genes, 389
Paracetamol, 830
Paragangliomas (PGLs), 753
ParagonDX/SmartCycler, 534
Parallel controlled pharmacogenetic diagnostic study, 318
Paroxetin, 101–102
Paroxetine, 7–9
Parthenogenetic embryonic stem (pES) cells, 209–210
Partial LCAT deficiency, 726
Passive reabsorption, 348
Pasteur, Louis, 933
Patatin-like phospholpase-3(*PNP*LA3), 865
Pathos, 200
PathVisio, 279–280
Patient-derived xenografts (PDX), 812
PCA analysis, 988
PCR enrichment, 139
Peak expiratory flow (PEF) measurements, 375, 622
Pediatric asthma, 622–623
Pediatric pharmacogenomics, 931–933
  benefit, risk, and uncertainty, ethical significance of, 935–943
    assent, 940–943
    minimal risk, 943
    “consent”, 938–940
  categories of pediatric research and requirements for IRB approval, 942
  federal regulations and national healthcare priorities, 943–945
  future implications of, 946–948
  genotypic approaches to therapy, 947
  historical perspective, 933–935
  translation of pharmacogenomics into clinical setting, 945–946
Pediatric Research Equity Act (PREA), 944
Pediatric Trials Network (PTN), 944
PEG-IFN/RBV, 12–13
Peginterferon alfa-2b, 7–9
Peginterferon and ribavirin therapy, 866–869
  *IL*28B genotyping, 866
    and pegIFN therapy, 867–869
    and peginterferon/ribavirin therapy in HCV-1, 866–867
    and rapid virological response (RVR), 867
    and short-treatment duration, 867
Penicillium griseofulvum, 879
Pentamidine, 794, 910–911
Peptide transporters (PEPTs), 890
Peptide-loading complex (PLC), 438
Peptides/sequencing MS, identification of, 158–159
Peptidyl arginine deiminase (PADI), 653–654
Percutaneous coronary intervention (PCI), 374
Peripheral blood mononuclear cells (PBMCs), 91–92
Peroxisome proliferator-activated receptor gamma (PPARγ) activator, 707, 1055
Peroxisome proliferator-activated receptor γ2 (PPAR-γ2), 1008
Peroxisome proliferator-activated receptor-α (PPARα), 394–395, 741
Peroxisome proliferator-activated receptor-γ co-activator 1β (*PGC*1β) gene, 390
Peroxisomes, 31
Perphenazine, 7–9, 12–13
Personalization, 801–802
Personalized medicine, 172, 933, 945
  of Alzheimer’s disease, 563–564
  for dementia and neurodegenerative disorders, 564–565
Personalized Medicine Program (PMP), 926
Pertuzumab, 7–9
PG4KDS pharmacogenetics implementation program, 927
P-glycoprotein (Pgp), 552, 888
  and anesthesia, 816
PGRN (Pharmacogenomics Research Network), 927
PGx clinical trials, 1021
  of antipsychotics in schizophrenic patients, 1030–1032
  in Brazil
    antiretrovirals, 1024–1025
    L-thyroxine in differentiated thyroid cancer, 1028
    NSAIDs, 1021–1024

    - warfarin-dosing algorithms, 1026–1028
  - of immunosuppressants in renal transplant patients, 1028–1031
- PGx Working Party (PgWP), 953
- PGx/biomarkers
  - clinical trial design and drug development utilizing, 960–963
    - cohort design, 960–961
    - designs of clinical trials, 960–962
    - enriched design, 961–962
    - stratified design, 961
  - confirmatory clinical trial with, 962
  - in drug labels, 960
  - exploratory clinical trials with, 962
  - future challenges, 963–964
  - PK evaluation with, 962
  - postapproval clinical trials with, 962–963
  - process, 957–958
    - case example, 958
  - qualification, 956–960
  - regulatory activities and guidelines relating to, 951–956
    - European Medicines Agency (EMA), 953–954
    - Pharmaceuticals and Medical Devices Agency (PMDA), 954–955
    - United States Food and Drug Administration, 955–956
  - significance, 956–957
  - for a specific drug product and information in the drug label, 958–960
  - study design and development strategy for, 962–963
- Pharmaceutical Research and Manufacturers of America (PhRMA), 944
- Pharmaceuticals and Medical Devices Agency (PMDA), 954–955, 957–959
- Pharmacodynamic pharmacogenetics, 373–379
  - adverse drug reaction (ADR), 378–379
    - malignant hyperthermia, 378–379
  - genetic variability that indirectly alters drug response, 376–378
    - factor V, prothrombin, and the oral contraceptive pill, 377–378
    - glucose-6-phosphate dehydrogenase (G6PD) deficiency, 377–378
  - and targets for drug action, 373–376
    - antiplatelet drugs, 373–376
    - statins, 373
- Pharmacodynamic pharmacogenomics, 5
  - of insulin resistance, 84–85
- Pharmacodynamics, 365–373
  - basic principles, 366
  - drug action, measurement of, 370–373
    - allosteric modulation, 372
    - competitive antagonism, 371–372
    - desensitization and tachyphylaxis, 372–373
    - dose–response relationship, 370–371
    - efficacy and potency, 372
    - non-receptor-mediated drug response, 373
    - therapeutic index, 371
  - historical aspects, 365–366
  - receptors, 366–370
    - binding, dose-response, and regulation, 369–370
    - enzymes, 369
    - G protein–coupled receptors (GPCRs), 368
    - ion channels, 368–369
    - mechanisms by which drugs act on, 366–368
    - nuclear receptors, 369
    - receptor protein kinases and guanylyl cyclases, 368
    - transport receptors, 369
- Pharmacogenetic (PGx) testing, 931–932, 937, 941, 946
  - in children, 933
  - future perspectives, 927–928
  - point-of-care, 924–925
  - regulation, 924
  - turnaround time, 924
- Pharmacogenetic biomarkers
  - and companion diagnostics, 13–19
  - evaluation, for clinical use, 17–19
- Pharmacogenetics, 5, 323
- Pharmacogenomic Resource for Enhanced Decisions in Care and Treatment (PREDICT), 926
- Pharmacogenomic studies
  - creating stem cell banks for, 219
  - generation of ethnically diverse pluripotent stem cells for, 218–219
- Pharmacogenomics, 5
  - genome-wide association studies in, 12–13
  - historical aspects of, 5
- Pharmacogenomics and stratified medicine, clinical trials in, 309
  - phase I studies, 309–312
  - phase II studies, 312–314
  - phase III studies, 314–319
    - randomized controlled trials (RCTs), 316–319
  - phase IV trials, 319
  - retrospective analysis of clinical datasets, 319
- Pharmacogenomics Knowledge Base (PharmGKB), 280, 985
- Pharmacogenomics Research Network (PGRN), 922–923, 944
- Pharmacokinetic pharmacogenomics, 341
  - ADME, 349–360
    - CYP450 enzymes, 350–354
    - drug transporters, 357–360
    - non-CYP450 enzymes, 354–357
  - principles of pharmacokinetics, 341–349
    - absorption, 341–342
    - clearance, 349
    - distribution, 342–343
    - excretion, 346–348
    - half-life, 348–349
    - metabolism, 343–346
    - zero- and first-order kinetics, 349
- PharmGKB, 289
  - clinical annotations, 294–297
  - clinical implementation, 301–303
  - of evidence for clinical annotations, 297
  - homepage, 290
  - Knowledge Pyramid, 291
  - projects in development, 303–304
  - variant annotation, 292
  - variant annotations, 291–294
  - VIPs and pathways, 297–301
  - website, 290–291
- Phase I enzymes, 816–818
  - CYP2C, 817–818
  - CYP2C19, 818
  - CYP2D, 817
  - CYP2D9, 817
  - CYP2E1, 817
  - CYP3A, 817
  - CYP3A5, 817
- Phase II enzymes, 819–820
  - glutathione S-transferases (GSTs), 819–820
  - N-acetyltransferases (NAT1 and NAT2), 819
  - UDP-glucuronosyltransferases (UGTs), 819–820
- Phenobarbital, 445–449
- Phenothiazines 794
- Phenotype, 235–239
  - assessment of genetic basis of trait, 237
  - familial aggregation, 237
  - heritability, 237–239
    - of pharmacogenomic traits, 239
  - intermediate phenotypes and endophenotypes, 236–237
  - recurrence risk, 237
  - standardized, 237
- Phenotype Standardization Project (PSP), 237, 787
- Phenotypic consistency, 75
- Phenprocoumon, 352
- Phenytoin, 7–9, 333, 352, 418, 445–449, 1055
- Pheochromocytomas (PCCs), 753
- Phosphorylation, 368
- Phytohemagglutinin (PHA), 91
- P-i model, 443–444
- PiggyBac, 224
- Pimozide, 7–9
- Pioglitazone, 691, 697, 701, 708, 1055
- Piroxicam, 1022
- Pitrikinra, 645
- PK/PD studies, 956
- Placenta-specific ATP binding cassette transporter (ABCP). *See* ABCG2
- Plasma, 150–151
- Plasma cholinesterase. *See* Butrylcholinesterase (BChE)
- Plasma-protein binding, 342–343
- Plasminogen activator inhibitor (PAI), 482
- *Plasmodium parasites*, 902, 905
- Platelet endothelial aggregation receptor-1 (*PE*AR1) gene, 526–527
- Platelet Function Analyzer 100 (PFA-100), 515
- Platelet turnover and reticulated platelets, 518
- Platelet–leukocyte interactions, 518
- Platelets, 509–511
- Platinum-based therapies, in pancreatic cancer, 809
- PLSDA, 163–164
- Pluripotent stem cells, 208, 220
  - epigenetic variation among, 215–216
  - epigenomics in

PMDA Omics Project (POP), 955
Pneumocystis jirovecii pneumonia (PJP), 412
Point-of-care pharmacogenetic testing, 924–925
  future perspectives, 927–928
*P*OLG, 473
Poly(A)-binding protein (PABP), 49–51, 63
Polyenes, 879–881
POPular, 515
Population diversity in Africa, 973–979
Population genetics, 239–247
  linkage disequilibrium, 241–243
    Hardy-Weinberg equilibrium, 242–243, 245–246
    HWE in association studies, 243
  linkage studies, 243–247
    identical by descent (IBD), 246–247
Population-based study designs, 249–253
  study power and coverage, 252–253
  tests of association, 250–252
Porphyria, 820–821
  anesthetic drugs and, 821
Posaconazole, 883–884, 891, 893–895
Positional cloning, 626–628
Positive predictive value (PPV), 867
Positron emission tomography (PET) tracers, 569
Post-translational modifications (PTMs), 63, 148
Post-transplant lymphoproliferative disease (PTLD), 90
Post-transplantation HCV, 868
Potassium channel, 788
Potency, 372
Prader-Willi syndrome (PWS), 65
Pragmatic RCTs (PRCTs), 316
Prasugrel, 7–9, 514
Pravastatin, 7–9, 416–417, 729
Preclinical genetically engineered mouse models (GEMM), 812
Precursor miRNA (pre-miRNA), 47–48, 385
PREDICT-1 prospective controlled trial, 419
Predictive biomarkers, 953–954
Preeclampsia, 171
Pre-emptive pharmacogenetic testing, 925–927
  1200 Patients Project, 926–927
  CLIPMERGE–PGx, 927
  future perspectives, 927–928
  Personalized Medicine Program (PMP), 926
  PGRN (Pharmacogenomics Research Network), 927
  Pharmacogenomic Resource for Enhanced Decisions in Care and Treatment (PREDICT), 926
  St. Jude PG4KDS pharmacogenetics implementation program, 927
Pregnane X receptor (PXR), 852
“Prescribing information”, 958
Primary dyslipidemias, 722–727
  primary HDL abnormalities, 725–726
  primary LDL abnormalities, 722–725
  primary triglyceride abnormalities, 726–727
Primary miRNAs (pri-miRNAs), 385
Primary monogenic dyslipidemias, 723–724
Primary transcripts (pri-miRNA), 47–48
Principal component analysis (PCA), 254, 980–981
Procainamide, 445–449
Prognosis prediction tools, 811–812
Prokaryotes, 29
Prokaryotic cells, 29–30
Prometaphase, 32
Propafenone, 7–9
Prophase, 32
Propofol, 821
Propranolol, 7–9
PRO-PredictRx TPMT®, 944–945
Proprotein convertase subtilisin/kexin type 9 (PCSK9), 733, 737–738
  deficiency, 725
Prostaglandin G2 (PGG2), 373
Protamine sulphate, 486
Protease inhibitor (PI), 1025
Protease-activated receptor (PAR)-1, 510
Protein, 99
  in urine, 149–150
Protein data bank file format, 271–272
Protein modeling, 272
Protein stability, 269–271
  allowed main-chain conformation, 270
  alpha helices, 270
  beta sheets, 270–271
  permitted torsion angles, 270
  turns and loops, 271
Proteomic-pattern analysis, 164–167
Proteomics, 147
  analytical platforms, 152–157
    CE-MS, 153, 155–156
    gel-based approaches, 152–155
    LC-MS, 153, 156
    MudPIT, 156–157
    non-gel-based approaches, 155–157
    SELDI-TOF MS, 157
  in biomarker research, 168–169
  data analysis, 163–167
    biomarker pattern analysis, 164–167
    data integration and OMICS analysis, 167
    downstream processing, 163–164
  discovery and validation, 167–172
    cancer markers, 167
    cardiovascular diseases, 167–170
    kidney diseases, 170–171
    neurological complexes, 171–172
    preeclampsia, 171
  ionization techniques, 157–158
    electrospray ionization (ESI), 157–158
    matrix-assisted laser desorption ionization (MALDI), 158
  peptides/sequencing MS, identification of, 158–159
  personalized medicine, 172
  quantitative approaches, 159–163
    isotope-coded affinity tag (ICAT) method, 159–160
    iTRAQ, 160
    label-free methods, 161
    multiple reaction monitoring (MRM), 161–163
    selected reaction monitoring (SRM), 161–163
    SILAC, 160–161
  sample types and sample preparation requirements, 148–152
    plasma, 150–151
    saliva, 150–151
    serum, 150–151
    tissue, 150–152
    urine, 149–150
  and systems biology analysis of systems perturbations, 169–170
  workflow, 149
Proteomics Identification Database (PRIDE), 279
Prothrombin complex concentrates (PCCs), 492
Prothrombin time (PT), 480–481, 490
Proton NMR, 187–188
Proton-pump inhibitors (PPIs), 10, 516–517
Protriptyline, 7–9
Proximal interphalangeal (PIP) joints, 654
Pseudocholinesterase. *See* Butrylcholinesterase (BChE)
Pseudohypoaldosteronism type II, 752–753
Public databases for biomedical research in humans, 38–39
Pulmonary arterial hypertension (PAH), 84
Pulmonary embolism (PE), 483
Pulmonary thrombosis/thromboembolism (PTE), 377
Purine pathway enzymes, 664–665
“Purkinje fibers”, 779–780
PyMOL, 271
Pyrazinamide (PZA), 7–9, 841, 420, 441–442

Q

Qpalma, 134
QT interval, 780–781
  corrected QT interval (QTc), 781
QT prolongation
  drugs associated with, 792–793, 795
  inducing drugs, 791–794
  mechanisms of proarrhythmia in, 788–790
QT-prolonging drugs, 795
Quantification of transcript abundance, 135
Quantitation, 198
Quantitative approaches, 159–163
  isotope-coded affinity tag (ICAT) method, 159–160
  iTRAQ, 153, 160
  label-free methods, 161
    design of MRM transitions/MRM assays, 162
    generation of an SRM method, 162
    label-free using ion intensities, 162–163
    label-free using spectral counting, 163
  multiple reaction monitoring (MRM), 161–163
  selected reaction monitoring (SRM), 161–163
  SILAC, 160–161
Quantitative multiplex PCR of short fluorescent fragment (QMPSF), 98
Quantitative NMR, 189

- Quantitative trait loci (QTL), 75, 81
- Quantitative trait mapping, 81
- Quetiapine, 12–13
- Quinidine, 7–9
- Quinine, 445–449, 794

R

- Rabeprazole, 7–9
- Rabies vaccine, 933
- Race, 973
- Radiofrequency (RF) pulses, 187
- Randomized clinical trials (RCTs), 3
- Randomized controlled phase II study, 313
- Randomized controlled trials (RCTs), 316–319
- Randomized discontinuation design, 313
- Randomized phase II selection design, 312–313
- Rapid FDA/CEIVD-approved clopidogrel pharmacogenetic testing, 529–534
- RAPID GENE trial, 924–925
- “Rapid sample-to-result” platform, 925
- RAPID STEMI, 924–925
- Rapid virological response (RVR), 867
- RAR-related orphan receptor A (*R*ORA) genes, 553–554
- Rasburicase, 7–9
- Rat Genome Database (RGD), 79
- Rat models, 74
- Read assembly, 130
  - *De N*ovo Assembly, 130
- Read quality control, 129–130
- Recombinant-inbred strains (RIS), 76
- Recurrence risk, calculation of, 238
- Rede Nacional de Farmacogenética. *See* Refargen
- Reduced repolarization reserve, 790–791
- Reduced-representation bisulphite sequencing (RRBS), 61
- Reduction, 80
- Reduction reactions, 345
- Refargen, 1015–1021
- Reference sequence, 133–134
- Refinement, 80
- Regenerative medicine, 207
  - epigenomics in pluripotent stem cells for, 216–218
- Relational database management system (RDBMS), 281
- Relative risk (RR), 252
- Remifentanil, 819
- Renal excretion, 346–348
- Renal papillary antigen-1 (RPA-1), 958
- Renin-angiotensin-aldosterone system (RAAS), 748, 756, 763
- Repaglinide, 690, 695
- Replacement, 80
- Repolarization, detecting drug-induced effects on, 795
- Research decisions, children’s participation in, 941
- Research ethics, historical perspective, 933–935
- Resectable pancreatic cancer, 811–812
  - adjuvant chemotherapy, 812
  - adjuvant radiotherapy, 812
  - neoadjuvant therapy versus surgery upfront, 811–812
- Retroviruses, 210–211
- Rett syndrome, 220
- Reverse cholesterol transport, 718, 720
- Rheonix CARD$^{®}$, 534
- Rheumatoid arthritis (RA), 653
  - biologic DMARDS, pharmacogenetics of, 668–674
    - rituximab, 673–674
    - tocilizumab, 674
    - tumor necrosis factor antagonists, 668–673
  - diagnosis, 654
  - epidemiology, 653
  - future directions, 674
  - immunologic features, 653–654
  - nonbiologic DMARDS, pharmacogenetics of, 655–668
    - azathioprine (AZA), 663–665
    - hydroxychloroquine (HCQ), 666–668
    - leflunomide, 661–663
    - methotrexate (MTX), 655–661
    - sulfasalazine (SSZ), 665–666
  - pathogenesis, 654
  - pharmacogenetics and stratified medicines in, 655
  - treatment, 654–655
- Rheumatoid factor (RF), 654
- Rhodesiense HAT, 906
- Ribavirin therapy, 866–869
- Ribavirin-induced hemolytic anemia
  - *I*TPA polymorphisms and, 864–865
- Ribosomal Database Project Classifier, 137–138
- Ribosomal RNAs (rRNAs), 42
- Ribosomes, 31
- Rifampicin (RMP), 7–9, 420, 841, 850, 852–853
- RISC loading complex (RLC), 48
- Risk stratification, proposed methods of, 802
- Risperidone, 7–9, 12–13
- Ritonavir, 1025
- Rituximab, 673–674
- Rivaroxaban *See also* Factor Xa inhibitors, 488, 492
- Rivastigmine, 582
- RNA interference (RNAi), 78–79
- RNA-induced silencing complex (RISC), 78–79, 385
- RNA-seq, 99–100, 134
- Roadmap Epigenomics Project, 38–39
- Roche and Life Technologies, 127
- Roche-454, 127
- Rodent models, 74–79
  - collaborative cross (CC), 75–76
  - congenic strains, 76–77
  - consomic strains, 77
  - heterogeneous stocks, 75
  - inbred strains, 75
  - mouse models, 74
  - outbred rodent stocks, 74–75
  - rat models, 74
  - recombinant-inbred strains (RIS), 76
  - transgenic rodent models, 77–79
    - Cre-LoxP Technology, 77
    - recent transgenic advances, 77–79
- Rodent models in pharmacogenomics, 82–85
  - anxiety disorders, 82
  - bipolar disorders, 82
  - pharmacodynamic/pharmacogenomic modeling of insulin resistance, 84–85
  - pharmacogenomics and PAH, 84
  - pharmacogenomics of pain and analgesia, 83–84
  - pharmacogenomics using xenograft mouse models of acute leukemia, 82–83
- Rosiglitazone, 691, 698, 708
- Rosuvastatin, 416–417
- Roussel Uclaf Causality Assessment Method (RUCAM), 471
- Royal Dutch Association for the Advancement of Pharmacy-Pharmacogenetics Working Group (KNMP-PWG), 922–923
- rs4149601, 765
- Ryanodine receptor 3 gene (*R*YR3), 770–771
- Ryanodine receptor gene (RYR1), 379
- Ryanodine receptor type 1 (RYR1), 414

S

- Saliva, 150–151
- Sanger, Fred, 125
- Sardilipin, 583–584
- Scavenger receptor class B1 (SR-B1), 718
- Schizophrenia (SCZD), 220
- Schizophrenic patients, PGX of antipsychotics in, 1030–1032
- Schmeideberg, Oswald, 366
- Screening, 801–802
- Scythe software, 130
- SCYX-7158, 913
- Second-generation sequencing (SGS), 126–129
  - applications, 128–129
  - clonal amplification, 127
  - DNA library preparation, 127
  - sample preparation, 128
  - sequencing DNA fragments in parallel, 127–128
  - spatial separation of fragments, 127
  - third-generation methods (TGS), 128
- Secreted Protein Acid and Rich in Cysteine (SPARC), 810
- Seed sequence, 386–387
- SELDI-TOF-MS biomarker model, 153, 157, 164, 167, 170–172
- Selected reaction monitoring (SRM), 153, 161–163
- Selective serotonin-reuptake inhibitors (SSRIs), 794, 924
- Selective target circularization methods, 139
- Separation and detection techniques, 152–157
- Sequence and read count, 131
- Sequence Read Archive (SRA), 276
- Sequencing, 133–134
  - DNA fragments in parallel, 127–128
  - reference sequence, 133–134
  - RNA-seq read-mapping process, 134
- SEQUEST, 147
- Serious drug-induced adverse events (SAEs), 467

Serotherapy, source of neutralizing antibodies for, 103
Serotonin (5-HT), 545–546
Serotonin 1A receptor (5-HT1A receptor), 546
Serotonin transporter (SERT), 545–546
Sertraline, 333
Serum, 150–151
Serum cholinesterase. *See* Butrylcholinesterase (BChE)
Serum resistance-associated (SRA) gene, 903
Severe asthma with fungal sensitization (SAFS), 621
Severe cutaneous adverse reactions (SCARs), 302–303, 440
Sevoflurane hepatotoxicity, 828
Sex cells, 29–30
Sex chromosomes, 33
Shared epitope (SE), 653
Short interfering RNAs (siRNAs), 78–79
Short interspersed repeats (SINES), 42
Short-acting β2 adrenoceptor agonists (SABA), 374–375
Side-chain modeling, 274–275
SILAC, 153, 160–161
Silylation, 197
Simvastatin, 12–13, 354, 416–417
Simvastatin and Ezetimibe in Aortic Stenosis (SEAS), 736
Single-gene studies, 5–12
  predicting drug dosage, 10
  predicting drug toxicity, 10–12
  predicting efficacy, 5–10
Single-ion monitoring (SIM), 196–197
Single-nucleotide polymorphism (SNP), 12, 38, 42, 53–54, 82, 97–98, 111–112, 239–241, 276–277, 323–324, 375, 395–396, 420, 439, 543, 657, 684, 845, 859–860, 888, 890, 985, 1001, 1010, 1037
  effect, on gene expression, 53
  linkage and haplotype blocks, 53–54
Single-organ drug hypersensitivity syndromes, 441–442
Sitosterolemia, 725
SJS/TEN, 441
SLBP (stem-loop–binding protein), 63
SLC transporters, 359–360
  organic anion transporters (OATs), 360
  organic anion–transporting polypeptides (OATPs), 359
  organic cation transporters (OCTs), 359–360
*SLC22*A1, 1050
*SLC22*A1 gene polymorphisms and metformin response, 687–702
SLC30A8, 1006
*SLCO*1A2 gene, 359
*SLCO*1B1 gene, 359, 416, 742, 1050
  and myopathy, 730
Sleeping sickness. *See* Human African trypanosomiasis (HAT)
Sleep-onset rapid-eye-movement (SOREM), 906
Small cytoplasmic ribonucleoproteins (scRNPs), 45
SN-38 glucuronidation, 12
S-nitrosoglutathione reductase (*GS*NOR), 639
*S*OD2 gene, 473
Sodium, effect on BP, 749–753
Sodium Benzoate, 7–9
Sodium bisulfite, 61
Sodium channel α-subunit type 1 (SCN1A), 1010–1011
Sodium Phenylacetate, 7–9
Sodium Phenylbutyrate, 7–9
Solute-carriers (SLCs), 342, 815–816
Sorafenib, 237
*SOR*T1 gene, 116
SOUND-C2, 871
Sparfloxacin, 794
Spartan RX system, 532–534
Spatial separation of fragments, 127
Spermatogenesis–associated, serine-rich 2-Like (*SPATS*2L), 639
Spinal muscular atrophy (SMA), 220
Spin-lattice relaxation time, 189
SpliceMap, 134
Splicing, 44
SQTL, identification of, 118
SREBP cleavage–activating protein (SCAP), 716–717
16SrRNA, 137–138
St. John's wort, 1011
St. Jude PG4KDS pharmacogenetics implementation program, 927
ST elevation myocardial infarction (STEMI), 509–510
Standard phase II design, 312
Statins, 373, 516, 728–729, 1054
  dosing, 729
  efficacy of, 731
  metabolism, 732
  pharmacodynamic pharmacogenomics of, 733–738
  pharmacokinetic pharmacogenomics of, 731–733
  and skeletal muscle toxicity, 415–417
Statistical HeterospectroscopY (SHY), 190–191
Statistical TOtal Correlation SpectroscopY (STOCSY), 190–191
  NMR hyphenation, 191
  water suppression, 191
Stem cell research using LCLs, 103–104
Stem cells, 207
  adult, 208
  alteration of gene expression in, 215
  for cancer therapy, 226
  chromatin modifications of, 214–215
    polycomb group (PcG) protein complexes, 214
    trithorax-(TrxG-) group protein complexes, 214–215
  culture methods, 211–212
  for disease modeling, 219–221
  epigenetic modifications of, 212–214
    DNA methylation, 212–213
    histone modifications, 213
    small interfering RNAs (siRNAs), 213–214
  epigenomics of, 215–218
    epigenetic variation among pluripotent stem cells, 215–216
    in pluripotent stem cells for regenerative medicine, 216–218
  fetal, 208–209
    amniotic fluid (AF) stem cells, 209
    amniotic membrane stem cells, 209
    embryonic stem (ES) cells, 209
    umbilical cord blood (UCB) stem cells, 209
    Wharton's jelly stem cells, 209
  future prospects, 226–227
  genomic imprinting in, 215
  germline-derived pluripotent stem (gPS) cells, 210
  Induced Pluripotent Stem (iPS) cells, 210–211
  parthenogenetic embryonic stem (pES) cells, 209–210
  pharmacogenomics of, 218–219
    creating stem cell banks, 219
    generation of ethnically diverse pluripotent, 218–219
    use of stem cells for toxicogenomic studies, 219
  for toxicity modeling, 221–226
    in cardiotoxicity studies, 221–223
    in hepatotoxicity studies, 223–224
    in neurotoxicity studies, 224–226
Sterol regulatory element–binding protein (SREBP) pathway, 716–717
Stevens Johnson syndrome (SJS), 417
Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN), 10
Stratification, 801–802
Stratified medicine, 14–15, 808–812
  for locally advanced pancreatic cancer, 811
  for metastatic pancreatic cancer, 808–811
  oral antidiabetic medicines, 709–710
  for resectable pancreatic cancer, 811–812
Streptomyces nodosus, 879
Stress-induced phosphoprotein 1 gene (*STI*P1), 644
STRUCTURE software, 980
Study of Heart and Renal Protection (SHARP), 735–736
Succinate dehydrogenase subunit B and D (SDHB, SDHD), 753
Succinylcholine, 819
  -induced neuromuscular blockade, 821
Sudden cardiac death (SCD), 789
Sufentanil, 816
Sulfamethoxazole (SMX), 412, 418, 442–443, 445–449
Sulfasalazine (SSZ), 665–666
  metabolism, 666
  *N*AT2, 665–666
  pharmacogenetics of, 666
Sulfonylureas, 699, 703–706
  disease-associated gene polymorphisms, 705–706
  inward rectifier K+ channel Kir6.2, 705
  sulfonylurea pharmacokinetics, 703–705
Sulfotransferases (SULT), 345
Sulphobutylether beta cyclodextrin sodium (SBECD), 883
Sulphonomides (sulfa antimicrobials), 445–449

Sulphotransferase 1A1 (SULT1A1), 1049
*SULT*1A1 gene, 396
Support vector machines (SVM), 163
Suramin, 910–911
Swiss-Prot, 279
Synovial membrane hyperplasia, 654
Systemic lupus erythematous (SLE), 663–664
Systolic blood pressure (SBP)
  APOE-related, 582–583
Systolic BP (SBP), 747

T

T allele, 396
T box 21 and neurokinin 2 receptor genes, 619
Tacrine, 582
Tacrolimus (TAC), 1028, 1055–1056
Tacrolimus, PGxplore images for, 304
TAILOR-PCI (Tailored Antiplatelet Therapy Following PCI), 925
Tamoxifen, 5–10, 395, 1052
Tangier disease. *See* Analphalipoproteinemia
TaqMan, 529, 532
TAR RNA binding protein (TRBP), 385
Tarceva®, 807
Tardive dyskinesia, 410
Targeted whole-exome, 985
TATA-binding protein (TBP), 46
Taxanes (paclitaxel), 357–358
T-cell activation, 654
T-cell receptor (TCR), 438
*TCF7*L2 gene, 684
Telaprevir, 7–9
Telophase, 32
Temazepam, 825
Terbinafine, 7–9, 885–886, 893, 895
Termination step of transcription, 45
Tet-assisted bisulphite sequencing (TAB-seq), 61
Tetrabenazine, 7–9
2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD), 3, 7–8, 394
Tetrahydrofolate (THF), 655
Tetramethylsilane (TMS), 197
Therapeutic drug monitoring (TDM), 313
Therapeutic index, 371
Thiazide diuretics, 756, 758–761
  alpha adducin (*A*DD1), 759
  *G*NB3, 759
  GWAS signals, 760–761
  *NED*D4L, 759–760
Thiazolidinedione (TZD), 707–709, 1008
  organic Anion–Transporting *Polypeptide* 1B1, 708
  pharmacodynamics, 708–709
  pharmacokinetics, 707–708
Thienopyridine, 511
Thioguanine, 7–9
6-Thioguanine, 345
Thiopentone, 820
Thiopurine, drug metabolism, 663
Thiopurine methyltransferase (TPMT), 10–11, 345, 355, 663–664, 937, 945–946, 1049, 1053
Thiopurines, 410–411
  adverse effects of, 410–411
  cost-effectiveness, 411
  perspectives of professional and regulatory bodies, 411
  TPMT genotype, 411
  TPMT phenotype, 411
Thioridazine, 7–9
Thiouric acid (TU), 663–664
Third-generation methods (TGS), 128
Threading, 275–276
Thrombin, 481, 510
Thrombin-activatable fibrinolysis inhibitor (TAFI), 481
Thromboxane A2 (TxA2), 373, 509–510
Thymidylate synthase (TYMS), 655
Thymine-DNA glycosylase (TDG), 58
Ticagrelor, 7–9, 514
Ticlopidine, 333, 422, 424
Tissue, 150–152
Tissue factor pathway inhibitor (TFPI), 481
TNF/IL-10, 667–668
TNRC6 protein, 49–50
Tocilizumab, 674
Tolbutamide, 701
Toll-like receptors (TLRs), 668, 865–866
Tolterodine, 7–9
Toluene, 389
Tophat maps, 134
Torsades de pointes (TdP), 469, 780–781, 796
  drug-induced (diTdP), 787, 791
  drugs associated with, 792–793
  -related pharmacodynamic interactions, 795
  risks of, 789, 794
Tositumomab, 7–9
Total body surface area (TBSA), 441
Total cholesterol, 719–721
TOtal Correlation SpectryoscopY (TOCSY), 189
Totipotent stem cells, 207
Toxic epidermal necrolysis (TEN), 417
Toxicity modeling, use of stem cells for, 221–226
  in cardiotoxicity studies, 221–223
  in hepatotoxicity studies, 223–224
  in neurotoxicity studies, 224–226
Toxicity of drug, predicting, 10–12
Toxicogenomic studies, use of stem cells for, 219
*T*PMT, 664
TPMT genotype, 411
TPMT phenotype, 411
Tramadol, 830
  and Acetaminophen, 7–9
Tramatenib, 18
Transcript reconstruction, 134–135
Transcription activator-like effector nucleases (TALENS), 78
Transcription factors (TFs), 278
Transcription start site (TSS), 115–116, 278
Transcriptomic applications, 133–135
  differential expression, 135
  library preparation, 133
  quantification of transcript abundance, 135
  sequencing, 133–134
    reference sequence, 133–134
    RNA-seq read-mapping process, 134
  transcript reconstruction, 134–135
Transcriptomic family of applications, 129
*T*rans-eQTL, 115
Transfer RNAs (tRNAs), 42, 60
Transformation efficiency, 90–92
Transgenic rodent models, 77–79
  Cre-LoxP Technology, 77
  recent transgenic advances, 77–79
Translational medicine, in China, 1012
Translational Pharmacogenetics Program, 927
Transmembrane proteins, 815
Transport receptors, 369
Transporter genes, 657–658
Transposable elements, 58
Trastuzumab, 7–9, 16–18
Treating to New Targets (TNT) trial, 721
Tree inference
  methods of, 267–269
  principles of, 266–267
Tree terminology, 265–266
Tretinoin, 7–9
Triangle association model, 101
Triazoles, 881–884
  fluconazole, 882
  itraconazole, 882–883
  posaconazole, 883–884
  voriconazole, 883
Tricyclic antidepressants (TCAs), 794
Trifluoroacetyl chloride, 827
Triglycerides, 721, 728, 740–741
Trimethylsilylpropionic acid (TMSP), 197
Trimipramine, 7–9
Triple-quadrupole mass spectrometer, multiple reaction monitoring experiment using, 194
Troglitazone, 422, 698
Troglitazone in the Prevention of Diabetes (TRIPOD), 709
True breeding, 35
Trypanosoma brucei, 901
  diagnosis, 908
  morphology of, 904
Trypanosoma brucei gambiense, 901–903, 905
  diagnosis, 907–909
Trypanosoma brucei rhodesiense, 901–903, 905
  diagnosis, 907
Trypanosoma congolense, 905–906
Trypanosoma cruzi, 903
Trypanosoma evansi, 903–904
Tryptophan hydroxylase (TPH), 547
T-tests, 164
Tuberculosis (TB), 835
  burden of, 836–838
    among children, 838
    incidence and prevalence, 836–838
    mortality, 838
  causative organism for, 836
  chemotherapy for, 840–842
    MDR and XDR TB, 842
    short-course, 841–842
    smear-positive TB, treatment of, 842
  in children, 844–845
    development and natural history, 844

    - diagnosis and treatment, 844–845
  - diabetes mellitus DM and, 843
  - diagnosis, 838–840
    - active disease and latent TB infection, 838–839
    - extrapulmonary TB (EPTB), 839–840
    - serodiagnostic tests, and imaging studies, 840
  - etiology and natural history, 835–836
  - impact of HIV/AIDS, 843, 849
  - pharmacogenomics and, 845–853
    - aminoglycosides, 853
    - drug-induced liver injury, predicting, 853
    - isoniazid, 845–852
    - rifampicin (RMP), 852–853
  - pharmacogenomics of, 845
    - genetic control of drug action, 845
  - risk factors, 842–844
  - TB incidence and HIV prevalence, 837
- Tumor necrosis factor (TNF) antagonists, 668–673
  - cytokines, 672
  - Fcγ receptor variants, 672
  - Genome-wide association studies (GWAS), 672–673
  - MHC gene polymorphisms, 671
  - MHC Microsatellites, 671
  - pharmacogenetics of, 669
  - RA risk variants, 671–672
  - signaling pathways, 672
  - TNF gene polymorphisms, 668–670
  - TNF receptor polymorphisms, 670–671
  - toll-like receptors, 672
- Tumor necrosis factor (TNF) polymorphisms, 852
- T-wave alternans, 789
- Two-dimensional COrrelation SpectroscopY (COSY), 190
- Two-dimensional NMR, 189–190
- Two-dimensional polyacrylamide electrophoresis (2D-PAGE), 154
- Type 1 cytokines, 905
- Type 2 diabetes mellitus (T2DM), 683, 1007–1008, 1010
- Type B adverse drug reactions, 414–424
  - drug induced liver injury (DILI), 421–424
    - amoxicillin clavulanate (AC), 422–423
    - flucloxacillin, 423
    - lumiracoxib, 423–424
    - ximelagatran, 423
  - drug-induced skin injury, 417–421
    - abacavir hypersensitivity, 417–419
    - allopurinol hypersensitivity, 420
    - antiepileptic drugs (AEDs), 421
    - carbamazepine hypersensitivity, 419–420
    - maculopapular exanthema reactions, 420–421
  - malignant hyperthermia and anesthesia, 414–415
    - genotype testing, 415
    - muscle contracture testing, 414–415
  - statins and skeletal muscle toxicity, 415–417

U

- UCS Transformation Format—8-bit (UTF-8), 281
- UDP Glucuronosyltransferases (UGTs), 345, 354–355, 819–820, 1028
- UDP polymorphisms, 852
- UGT1A1, 114–115
- UGT1A1*28 polymorphism, 12
- *UGT*1A6, 473
- UGT1A7, 1048
- Ultrafast 2D NMR (UF-NMR), 190
- Umbilical cord blood (UBC) stem cells, 209, 225
- Umbilical cord matrix stem cells (UCMSCs), 209
- Umbilical cord perivascular cells (UCPVs), 209
- Umbilical cord stromal cells (UCSCs), 209
- Unfractionated heparin (UFH), 484–486
  - adverse effects, 485–486
    - heparin-induced thrombocytopenia, 485–486
    - osteoporosis, 486
  - monitoring, 485
  - pharmacogenomics, 486
  - pharmacokinetics, 485
  - reversal, 486
  - structure and mechanism of action, 484–485
- UniParc, 279
- Unipotent stem cells, 208
- UniProt, 279
- United States Food and Drug Administration, 955–957
- Universal Protein Resource (UniProt), 279
- University of Chicago, 926–927
- University of Florida and Shands Hospital, 926
- 3’-Untranslated regions (UTRs), 44, 112–113, 386
- Unweighted pair group method with arithmetic means (UPGMA), 269
- Urban versus rural, 977
- Uridine diphosphate glucuronosyltransferase 1A1 (UGT1A1), 12, 113–115, 412, 1048
- Uridine diphosphate glucuronosyltransferases (UGTs), 412
- Uridinyl diphosphate glucotransferases (UGTs), 472
- Urine, 149–150
- Uronyl 2-sulphotransferase (UST) genes, 554
- US PharmacoGenetics Research Network (PGRN), 985

V

- Validation and replication, 255
- Valproate, 422
- Valproic Acid, 7–9
- Vanderbilt University, 425
- Vanderbilt University Medical Center, 926
- Variable importance in the projection (VIP), 163–164
- Variant calling applications, 131–133
  - example, 132–133
- Variant surface glycoproteins (VSGs), 906
- Vascular endothelial growth factor (*V*EGF), 551
  - signaling pathway, 237
- Vascular mechanisms
  - in hypertension, 753–754
- Vascular volume
  - effect on BP, 749–753
- Vasoactive intestinal peptide (VIP), 84
- Vasodilator-stimulator phosphoprotein (VASP) assay, 515, 517
- Vemurafenib, 7–9, 18
- Venous thromboembolism (VTE), 482–483
  - epidemiology, 483
  - management, 483
  - pathophysiology, 483
- Ventricular tachycardia (VT), 781, 791
- Verigene system, 529–532
- Vinblastine sulfate, 357–358
- VIPs and pathways, 297–301
- Visual analog scale (VAS), 657
- Vitamin K (phytomenadione), 490
- Vitamin K antagonists (VKAs), 489–491
  - interactions, 489–490
  - management of bleeding and reversal, 490–491
  - mechanism of action, 489
  - monitoring the anticoagulant effect of warfarin, 490
  - nonhaemorrhagic adverse effects, 491
  - pharmacokinetics, 489
- Vitamin K cycle, 498
- Vitamin K epoxide reductase (VKOR), 413, 498
- Vitamin K epoxide reductase complex 1 (VKORC1), 10, 498, 500–501
- Vitamin K supplements, 500
- *VKO*RC1 gene, 114–115, 413–414, 992, 1018–1019, 1051
- *VKO*RC1 genotype, 10
- *VKO*RC1 polymorphisms, 301
- VMATs (vesicular monoamine transporters), 548
- Volume of distribution (Vd), 343
- Voluntariness, 938
- Voluntary exploratory data submissions (VXDS), 955
- Voluntary genomic data submissions (VGDSs), 955
- von Willebrand Factor (vWF) and ADAMTS-13, 518
- Voriconazole, 7–9, 333, 883, 891, 894–895

W

- Warfarin (Coumadin), 7–9, 12–13, 352, 413–414, 497–498, 817–818, 1053–1054
  - application of pharmacogenomics to, 503–504
  - bleeding risk, 413–414
  - dose requirement
    - genes with major effects on, 500–502
    - nongenetic factors affecting, 499–500
  - future research directions, 504

  pharmacodynamics, 413
  pharmacogenomic factors affecting, 502–503
    calumenin, 502–503
    CYP4F2, 502
  pharmacokinetics, 413
  pharmacology of, 498–499
Warfarin pharmacogenomics, 10
Warfarin-dosing algorithms, 1026–1028
Water suppression, 191
WATER suppression by GrAdient-Tailored Excitation (WATERGATE), 191
Water suppression Enhanced by T1 (WET), 191
Waterfall model. *See* MacFarlane cascade
Wepfer, Johann Jakob, 366
West of Scotland Coronary Prevention Study (WOSCOPS) trial, 720
Wharton's jelly stem cells, 209
Whole genome association (WGA), 472
Whole-genome sequencing (WGS), 947–948, 985
Wilcoxon-Mann-Whitney tests, 164

X

Xanthine oxidase (XO), 663–664
Xenograft mouse models, pharmacogenomics using, 82–83
Ximelagatran, 12–13, 422–423, 445–449
XML tags, 281
XTAG P450-2C19 assay, 534

Y

Yoruba from Nigeria (YRI), 98

Z

Zebrafish Neurophenome Database (ZNP), 79
Zero- and first-order kinetics, 349
Zinc-finger nuclease (ZFN) technology, 77–78, 224
Ziprasidone, 12–13
Zonisamide, 445–449

This page is intentionally left blank